

CRANFIELD UNIVERSITY

Air Transport Management MSc

Predicting Passenger Willingness to Pay for
Ancillary Services on the BCN-LGW Route:
A Machine Learning Approach Using
Likert-Scale Survey Data



Faculty of Engineering and Applied Sciences

Student: Ariadna Sorribas Bono (s446330)

Supervisor: Berkcan Uyan

Word Count: 15,891

Document: MSc Thesis

Date set: August 10, 2025

Date due: August 11, 2025

Abstract

The growth of ancillary services has become a central driver of airline profitability, yet understanding passengers' willingness to pay (WTP) for these offerings remains a challenge. This thesis investigates the key factors shaping WTP for ancillary services among travellers on the high-traffic Barcelona (BCN)–London Gatwick (LGW) short-haul route and evaluates the potential of predictive modelling to forecast individual passenger WTP.

Deploying a structured survey (N=171), the study collects and analyses data on passenger demographics, travel behaviour, and attitudinal responses toward four key ancillaries: seat selection, checked baggage, in-flight meals, and Wi-Fi. A preprocessing pipeline encodes Likert-scale responses and constructs composite WTP scores, which serve as targets for a range of machine learning algorithms (Logistic Regression, Random Forest, Support Vector Machine, and XGBoost). Model performance is evaluated using metrics such as accuracy, F1 Score, and ROC-AUC, with further interpretability gained through SHAP values and permutation importance analysis.

The results reveal pronounced heterogeneity in WTP: seat selection and checked baggage appeal broadly, while in-flight meals show lower and dispersed WTP, and Wi-Fi has sharply polarised demand. Statistical and modelling analysis show that travel purpose, annual income, and age are the most significant predictors of ancillary WTP, while gender has negligible influence. Predictive models achieve strong performance, especially for passengers with clearly high or low WTP; however, the 'medium' WTP segment remains challenging to classify, suggesting opportunities for richer data and model refinement.

These findings highlight the potential of dynamic, segment-driven ancillary pricing and marketing. The study calls for strategic product design, targeted offers, and investment in predictive analytics, allowing airlines to align their ancillary strategies with evolving passenger needs. Limitations include sampling constraints and reliance on self-reported data; and future research should incorporate revealed-preference data, longitudinal analysis, and broader feature sets to further advance ancillary revenue management.

Acknowledgements

I would like to sincerely thank my supervisor for their guidance, knowledge, and encouragement at every step of this research. I am also very grateful to my colleagues and peers for their helpful feedback and support. A special thanks goes to all the survey participants, because without their contributions, this work would not have been possible. And above all, I am truly thankful for the constant support of my family, friends, and partner, whose patience and understanding helped me through this entire process.

“If we knew what we were doing, it would not be called research, would it?”

— **Albert Einstein**

Table of Contents

1	Introduction	1
1.1	Background and Motivation	1
1.2	Research Problem	3
1.3	Research Aim and Objectives	4
1.4	Scope and Structure	5
2	Literature Review	8
2.1	Ancillary Services in Aviation	8
2.1.1	Defining Ancillary Services	8
2.1.2	Evolution and Growth of Ancillary Services	9
2.1.3	Economic and Strategic Importance	9
2.1.4	Key Studies and Findings	10
2.1.5	Challenges and Criticisms	12
2.1.6	Future Directions	12
2.2	Willingness to Pay in Aviation	13
2.2.1	Defining Willingness to Pay in Aviation	13
2.2.2	Importance of WTP in Airline Revenue Management	13
2.2.3	Determinants of Willingness to Pay	14
2.2.4	Key Studies and Findings	14
2.2.5	Challenges and Criticisms	16
2.2.6	Future Directions	16

2.3	Survey-Based WTP Measurement	17
2.3.1	Survey Design and Methodology	17
2.3.2	Implementation in Aviation Contexts	18
2.3.3	Data Analysis and Interpretation	18
2.3.4	Key Studies and Findings	19
2.4	ML for Consumer Behaviour Prediction	21
2.4.1	Machine Learning Approaches in Aviation	21
2.4.2	Data Sources and Feature Engineering	22
2.4.3	Model Evaluation and Interpretability	22
2.4.4	Key Studies and Findings	22
2.5	Gap Analysis	24
3	Methodology	27
3.1	Research Design	27
3.2	Survey Design	29
3.3	Data Collection	32
3.4	Data Preprocessing	33
3.5	Machine Learning Models	34
3.5.1	Justification for Model Choices	34
3.5.2	Theoretical Background	35
3.6	Evaluation Metrics	37
3.6.1	Rationale for Metric Selection	38
3.6.2	Confusion Matrix	39
3.7	Feature Importance Analysis Methods	40
3.7.1	SHAP Values	40

3.7.2	Permutation Importance	41
3.8	Software and Tools	42
4	Results	44
4.1	Descriptive Statistics	44
4.1.1	Sample Composition	44
4.1.2	WTP for Ancillary Services	49
4.1.3	WTP by Passenger Characteristics	51
4.2	Model Performance	55
4.2.1	Performance Metrics by Ancillary Service	56
4.2.2	Summary of Best Models	58
4.2.3	Error Analysis: Confusion Matrices	61
4.3	Feature Importance Analysis	64
5	Conclution	72
5.1	Interpretation of Findings	72
5.2	Implications for Airlines	74
5.3	Limitations	75
5.4	Future Research Directions	76
	Bibliography	78
A	CURES Approval Letter	88
B	Survey Document	90
C	Data Analysis Workflow Pseudocode	93

List of Figures

4.1	Distribution of respondents by gender.	46
4.2	Distribution of respondents by age.	46
4.3	Distribution of respondents by annual income bracket.	47
4.4	Distribution of respondents by usual travel purpose.	48
4.5	Distribution of respondents by annual flight frequency.	48
4.6	Histogram and boxplot of WTP for seat selection.	49
4.7	Histogram and boxplot of WTP for checked baggage.	50
4.8	Histogram and boxplot of WTP for in-flight meals.	50
4.9	Histogram and boxplot of WTP for Wi-Fi.	51
4.10	WTP for each ancillary by gender (boxplot).	52
4.11	WTP for each ancillary by age (violin plot).	53
4.12	Mean WTP for each ancillary service by usual travel purpose (bar plot).	54
4.13	Mean WTP for each ancillary service by annual flight frequency (bar plot).	54
4.14	Heatmap of macro-averaged F1 scores for each model (columns) and ancillary service (rows).	60
4.15	Confusion matrix for the XGBoost model on seat selection.	61
4.16	Confusion matrix for the XGBoost model on checked baggage.	62
4.17	Confusion matrix for the SVM model on in-flight meals.	63
4.18	Confusion matrix for the Logistic Regression model on Wi-Fi.	63
4.19	SHAP plot of feature impact on predicting seat selection WTP.	65

4.20	Permutation importance of passenger traits for seat selection WTP. . . .	65
4.21	SHAP plot of feature impact on predicting checked baggage WTP.	66
4.22	Permutation importance of passenger traits for checked baggage WTP. . .	66
4.23	SHAP plot of feature impact on predicting in-flight meals WTP.	67
4.24	Permutation importance of passenger traits for in-flight meals WTP. . . .	68
4.25	SHAP plot of feature impact on predicting Wi-Fi WTP.	69
4.26	Permutation importance of passenger traits for Wi-Fi WTP.	69
4.27	Permutation importance of passenger traits for all ancillary services. . . .	70
A.1	CURES approval letter.	89
B.1	Page 1 of the passenger survey.	91
B.2	Page 2 of the passenger survey.	92
C.1	Visual flowchart of the data analysis workflow.	96

List of Tables

2.1	Key studies and findings on ancillary services in aviation.	11
2.2	Key studies and findings on WTP in aviation.	15
2.3	Key studies and findings on survey-based WTP measurment.	20
2.4	Key studies and findings on ML for consumer behaviour.	24
3.1	Example confusion matrix for a multi-class classification problem with Low, Medium, and High categories.	39
4.1	Sample distribution by demographic and travel behaviour variables. . . .	45
4.2	Performance metrics for each model predicting WTP for seat selection. The highest value for each metric is highlighted.	56
4.3	Performance metrics for each model predicting WTP for checked baggage	57
4.4	Performance metrics for each model predicting WTP for in-flight meals .	57
4.5	Performance metrics for each model predicting WTP for Wi-Fi	58
4.6	Summary of the best-performing model for each ancillary service based on macro F1 score, including the main performance metrics.	59

List of Acronyms

Acronym	Meaning
AI	Artificial Intelligence
API	Application Programming Interface
AUC	Area Under the ROC Curve
BCN	Barcelona El Prat Airport
BCN-LGW	Barcelona–London Gatwick (airport/route)
CB	Checked Baggage
CAA	Civil Aviation Authority (UK)
CNN-LSTM	Convolutional Neural Network
CV	Contingent Valuation
EU	European Union
FN	False Negatives
FP	False Positives
IM	In-Flight Meals
INE	Instituto Nacional de Estadística (Spanish National Statistics Institute)
KNN	K-Nearest Neighbour
LCC(s)	Low-Cost Carrier(s)
LGW	London Gatwick Airport
MAE	Mean Absolute Error
ML	Machine Learning
MLP	Multilayer Perceptron
NumPy	Numerical Python (Library)
OAG	Official Airline Guide

Acronym	Meaning
Pandas	Python Data Analysis Library
RF	Random Forest
ROC	Receiver Operating Characteristic (curve)
ROC-AUC	Receiver Operating Characteristic – Area Under Curve
SAF	Sustainable Aviation Fuel
SHAP	SHapley Additive exPlanations
SP	Stated Preference
SS	Seat Selection
SVM	Support Vector Machine
TN	True Negatives
TP	True Positives
UK	United Kingdom
UK Gov	United Kingdom Government
USD	United States Dollar
VS Code	Visual Studio Code
WF	Wi-Fi
WTP	Willingness To Pay
XGBoost	Extreme Gradient Boosting

Chapter 1

Introduction

1.1 Background and Motivation

The airline industry operates in a particularly challenging environment, characterised by intense competition, high fixed costs, and thin profit margins. Despite generating nearly \$1 trillion in annual revenues, airlines typically achieve net profit margins of only 2–4%, a fraction of what is seen in many other industries ([IATA, 2025](#)). Major expenses (including volatile fuel prices, labour costs, aircraft acquisition and maintenance, and airport fees) consume the vast majority of airline income, leaving little room for substantial profitability. Also, external factors such as economic cycles, regulatory requirements, and global events further increase financial pressures and volatility in the sector ([Singh, 2025](#)). Even in periods of strong demand, airlines often struggle to generate sustainable returns, making the search for new and resilient income sources a constant priority.

In recent years, ancillary revenue has become a key component of the airline industry's financial strategy. Traditionally, airlines relied primarily on ticket sales for revenue generation. However, increasing competition, fluctuating fuel prices, and evolving consumer expectations have driven carriers to diversify their income streams. Ancillary revenue (consisting of non-ticket sources such as baggage fees, seat selection, in-flight meals, and priority boarding) has grown rapidly, reaching an estimated \$148.4 billion worldwide in 2024, a significant increase from \$109.5 billion in 2019 ([Wilson, 2025](#)). This growth reflects a broader industry trend where ancillary services now account for a significant share of total airline revenue, with some low-cost carriers generating over 50% of their income from these sources ([Tathagat and Montanari, 2024](#)).

Moreover, the strategic importance of ancillary revenue is shown by its role in increasing profitability and mitigating financial risks. Airlines that successfully encourage passengers to spend more on ancillary services tend to achieve higher returns on invested capital compared to those with lower ancillary sales. This shift has also allowed airlines to offer lower base fares, making air travel more accessible while maintaining financial stability through supplementary offerings ([OAG, 2023](#)).

Given this growing reliance on ancillary services, understanding what drives passengers' willingness to pay (WTP) for these offerings is increasingly vital. WTP is influenced by a range of factors, including demographic characteristics, travel purpose, and perceived value of the service. Accurate understandings of WTP allow airlines to tailor their pricing strategies, optimise product bundles, and improve customer satisfaction, ultimately leading to increased revenue and competitive advantage ([Jander, 2024](#)).

On high-traffic routes such as Barcelona (BCN) to London Gatwick (LGW), the stakes are particularly high. This route is one of the busiest in Europe, serving a diverse mix of business and leisure travellers, and is operated by multiple carriers with varying service models ([AGN, 2024](#)). The competitive environment intensifies the need for airlines to differentiate themselves not only through base fares but also through the effective design and pricing of ancillary services.

Therefore, by systematically studying passenger WTP on the BCN-LGW route, airlines can better understand the unique preferences and sensitivities of this market segment. Such knowledge supports the development of targeted marketing strategies and personalised service offerings, which are essential for maximising ancillary revenue and improving the overall passenger experience. In this context, research into WTP is not only academically relevant but also useful for industry stakeholders trying to navigate the rapidly evolving commercial aviation sector.

1.2 Research Problem

Despite the increasing relevance of ancillary revenue in the airline industry, there remains a notable lack of clarity regarding the specific determinants that shape passenger willingness to pay for these services, especially within the context of individual air routes. While industry trends highlight the financial impact of ancillary offerings, the fundamental behavioural and demographic factors influencing passenger decisions are not yet fully understood. This gap is particularly pronounced on routes such as Barcelona (BCN) to London Gatwick (LGW), where passenger profiles are heterogeneous and market conditions are dynamic.

Airlines operating on the BCN-LGW corridor are confronted with the challenge of serving to a diverse customer base, each segment exhibiting distinct preferences and sensitivities toward ancillary services. The complexity of these consumer behaviours presents significant challenges for revenue management teams aiming to optimise service design and pricing. Without strong, data-driven insights, efforts to personalise ancillary offerings or to implement dynamic pricing strategies risk being ineffective or inconsistent with actual passenger expectations.

Moreover, the abundance of available data (ranging from transactional records to survey responses) presents both an opportunity and a challenge. Extracting useful information from such data requires advanced analytical approaches which can capture the multidimensional nature of passenger decision-making. Traditional methods may fall short in detecting subtle patterns or addressing the interaction of demographic factors and travel behaviours.

Given this context, the central research question addressed in this thesis is:

Which factors most significantly influence passenger willingness to pay for ancillary services on the BCN-LGW route, and to what extent can predictive models accurately forecast this willingness based on observed passenger characteristics and behaviours?

This question is motivated by the need to close the existing knowledge gap between industry practice and academic research. By investigating the predictors of willingness to pay, the study intends to provide airlines with evidence-based information for segmenting their customer base, refining ancillary service portfolios, and improving the overall passenger experience.

1.3 Research Aim and Objectives

To address the research problem identified, the following aim has been defined:

Research Aim

To predict passenger willingness to pay (WTP) for ancillary services on the Barcelona (BCN) to London Gatwick (LGW) route using machine learning techniques.

To achieve this aim, the research will target the following objectives:

1. **Design and administer a Likert-scale survey** to collect data on passenger demographics, travel behaviour, and willingness to pay for specific ancillary services on the BCN-LGW route.
2. **Preprocess and encode survey responses** by numerically transforming Likert-scale data and constructing composite WTP scores for each ancillary service, ensuring the dataset is suitable for machine learning analysis.

3. **Develop and evaluate predictive machine learning models** (including XGBoost, Random Forest, Logistic Regression, and Support Vector Machine (SVM)) to forecast passenger WTP for ancillary services. Model performance will be assessed using metrics such as accuracy, F1-score, and ROC-AUC.
4. **Interpret model results using feature importance techniques** (e.g., SHAP, permutation importance) to identify and quantify the key demographic and travel behaviour predictors that most influence passenger WTP.

1.4 Scope and Structure

To ensure that the research remains focused, this thesis sets clear boundaries for what will and will not be examined. The scope of the study is outlined below:

- **What is included:**

- Analysis is based on passengers travelling on the BCN-LGW route, which is used as a representative case study for short-haul European routes, and conclusions are intended to be relevant for similar short-haul contexts.
- Only ancillary services explicitly covered in the survey (seat selection, extra baggage, in-flight meals and Wi-Fi access) are considered.
- Primary data is collected through a structured Likert-scale survey, focusing on demographic information, travel behaviour, and stated WTP for the specified ancillary services.
- The research applies machine learning models (XGBoost, Random Forest, Logistic Regression, SVM) to predict WTP and uses feature importance for interpretation and segmentation.
- The study evaluates model performance using established metrics (accuracy, F1-score, ROC-AUC) and interprets results in the context of the selected route and services.

- **What is not included:**

- The research does not analyse medium- or long-haul routes, nor does it directly examine city pairs other than BCN-LGW; findings are not intended to be generalised beyond the short-haul segment.
- Ancillary services not included in the survey (e.g., in-flight entertainment, lounge access) are outside the scope.
- The study does not use transactional or purchase data, all analysis is based on self-reported survey responses.
- The focus is on predictive modelling and segmentation, not on establishing causal relationships or testing the impact of interventions.
- Broader airline revenue management strategies, regulatory considerations, or macroeconomic factors are not addressed.

The remainder of this thesis is organised to guide the reader through the research process, from the review of relevant literature to the presentation and discussion of findings. Each chapter is briefly described below:

- **Chapter 2: Literature Review**

Reviews prior research on ancillary services and WTP in aviation, survey-based measurement methods, machine learning applications in consumer behaviour, and identifies existing gaps.

- **Chapter 3: Methodology**

Details the survey design, data collection process, preprocessing steps, the machine learning models applied, segmentation and cluster analysis techniques, model validation, and software tools used.

- **Chapter 4: Results**

Presents descriptive statistics, model performance metrics, feature importance analyses, and findings from passenger segmentation.

- **Chapter 5: Discussion**

Interprets the results, discusses implications for airlines, outlines study limitations, and suggests opportunities for future research.

Chapter 2

Literature Review

2.1 Ancillary Services in Aviation

The aviation industry has evolved considerably over the last two decades, with ancillary services developing as an essential component of airline business strategies. These services, which include a variety of offerings in addition to the basic transportation of passengers, have redefined the airline value proposition, reshaped revenue models, and transformed the passenger experience ([OAG, 2023](#)). This subsection reviews the evolution, classification, and significance of ancillary services in aviation, considering key studies and recent industry developments.

2.1.1 Defining Ancillary Services

Ancillary services in aviation refer to supplementary products and services offered by airlines in addition to the core flight ticket. These include, but are not limited to: ([Liu et al., 2024](#))

- Checked baggage and carry-on fees
- Preferred or extra-legroom seating
- Priority boarding and check-in
- In-flight meals, snacks, and beverages
- In-flight entertainment and Wi-Fi
- Lounge access
- Travel insurance, car rentals, and hotel bookings

The term “ancillary” refers to offerings that are secondary or supplementary to the main service of passenger transportation. While historically many of these services were bundled into the ticket price, the trend toward unbundling has allowed airlines to offer a more customised travel experience (TTS, 2025).

2.1.2 Evolution and Growth of Ancillary Services

The emergence of ancillary services as a major income source can be traced to the deregulation of the airline industry and the subsequent rise of low-cost carriers (LCCs) in the 1990s and 2000s (Tathagat and Montanari, 2024). LCCs, such as Ryanair and Southwest, pioneered the unbundling of fares, charging separately for services that were once included in the base ticket price. This strategy was soon adopted by traditional full-service carriers looking to remain competitive and profitable in an increasingly challenging market environment (Sheridan Stavac and Giannetti, 2024).

Over the past decade, ancillary revenues have surged. In 2013, global ancillary revenue was estimated at \$42.6 billion; by 2023, this figure had grown to approximately \$118 billion, representing nearly a fourfold increase. For some airlines, particularly LCCs, ancillary services now account for over 50% of total revenue (Tathagat and Montanari, 2024). This shift emphasises the critical role of ancillaries in maintaining financial resilience in the context of volatile fuel prices, fluctuating demand, and competitive pressures.

2.1.3 Economic and Strategic Importance

Ancillary services offer a dual benefit for airlines: they diversify revenue streams and increase profit margins, while also providing passengers with greater choice and personalisation (Liu et al., 2024). In an industry characterised by thin margins, ancillaries provide a buffer against the volatility of ticket sales and external shocks (e.g., economic downturns, pandemics).

Relevant studies highlight several strategic advantages:

- **Revenue Stability:** Ancillaries help airlines survive fluctuations in base fares and demand cycles ([OAG, 2023](#)).
- **Profitability:** Airlines whose passengers spend more on ancillaries achieve higher returns on invested capital (ROIC) compared to those with lower ancillary spend ([Tathagat and Montanari, 2024](#)).
- **Customer Experience:** By unbundling services, airlines can accommodate to diverse passenger preferences, allowing travellers to pay only for services they value ([TTS, 2025](#)).
- **Competitive Differentiation:** Innovative ancillary offerings can be a differentiator in crowded markets, increasing brand loyalty and repeat business ([Sheridan Stavac and Giannetti, 2024](#)).

2.1.4 Key Studies and Findings

Recent studies on ancillary services in aviation collectively highlight the growing strategic importance of these offerings, each providing a different perspective yet linked by their common focus on revenue optimisation and industry evolution. [Liu et al. \(2024\)](#) synthesise and review the literature, emphasising the shift toward personalised and WTP-driven pricing models. [Lu \(2019\)](#) builds on this by applying quantitative revenue management methods to explore how tailored ancillary pricing benefits airline profitability. Complementing these, [Shaw et al. \(2021\)](#) take an empirical approach, focusing specifically on third-party ancillary partnerships, thus extending the analysis to new business model dimensions.

Author and Paper	Methodologies	Findings
Liu et al. (2024) <i>“A review of ancillary services in the airline industry”</i>	<ul style="list-style-type: none"> - Systematic literature review - Thematic synthesis 	<ul style="list-style-type: none"> - Ancillary revenues reached \$117.5 billion globally in 2023, up from \$31.5 billion in 2013 - WTP varies by segment (e.g., long-haul vs. short-haul, business vs. leisure) - Calls for AI-driven personalised pricing
Lu (2019) <i>“Impacts of ancillary services on airline revenue management”</i>	<ul style="list-style-type: none"> - Quantitative modelling - Scenario simulations 	<ul style="list-style-type: none"> - Ancillary services increased total airline revenue by up to 7–9% when dynamically priced - Segment-specific pricing outperformed uniform pricing by 11% in revenue uplift - Suggests bundling increases conversion
Shaw et al. (2021) <i>“Third party ancillary revenues in the airline sector: An exploratory study”</i>	<ul style="list-style-type: none"> - Empirical analysis - Case study review - Market trend mapping 	<ul style="list-style-type: none"> - Third-party ancillaries can account for up to 30% of total ancillary revenue - 70% of airlines surveyed partner with external providers for car rental and accommodation - Market shifting toward integrated ecosystems

Table 2.1. Key studies and findings on ancillary services in aviation.

2.1.5 Challenges and Criticisms

Despite their benefits, ancillary services are not without controversy. Important challenges include:

- **Customer Perception:** Some passengers view ancillaries as hidden fees or unfair charges, potentially decreasing their satisfaction and loyalty ([Sheridan Stavac and Giannetti, 2024](#)).
- **Complexity:** Managing a wide array of ancillary products requires sophisticated IT systems and efficient integration across sales channels ([OAG, 2023](#)).
- **Regulatory Scrutiny:** Authorities in some jurisdictions have enforced guidelines to ensure transparency and prevent fraudulent pricing practices ([Liu et al., 2024](#)).

2.1.6 Future Directions

The future of ancillary services in aviation is closely tied to developments in technology, data analytics, and personalised marketing. It is anticipated that airlines will improve their ancillary initiatives by: ([Tathagat and Montanari, 2024](#))

- Employing artificial intelligence and machine learning to predict passenger preferences and optimise offers in real time.
- Expanding digital channels for ancillary sales, including mobile apps and direct distribution platforms.
- Developing new ancillary products, such as carbon offsetting and wellness services, to meet evolving traveller expectations.

2.2 Willingness to Pay in Aviation

Willingness to pay (WTP) has become a key concept in the aviation sector, influencing how airlines price both core and ancillary services. As the industry evolves towards more personalised and dynamic pricing models, understanding the determinants and implications of passenger WTP is essential for maximising revenue, improving customer satisfaction, and guiding strategic decisions. This subsection explores the definition, measurement, and significance of WTP in aviation, using both foundational studies and recent innovations.

2.2.1 Defining Willingness to Pay in Aviation

WTP in aviation refers to the maximum price a passenger is prepared to pay for a specific air travel product or service, whether it be a flight ticket, seat upgrade, baggage allowance, or an ancillary offering. Unlike traditional fare structures, which segment passengers by rigid booking classes, WTP-based approaches aim to account for each traveller's price preferences and sensitivities. This shift is essential to the continuous transformation from class-based revenue management to dynamic, offer-optimised retailing ([Jander, 2024](#)).

2.2.2 Importance of WTP in Airline Revenue Management

Integrating WTP into revenue management systems allows airlines to better align their prices with what individual passengers are willing to pay. This alignment enables airlines to maximise revenue by capturing consumer surplus, while also reducing the risk of pricing out potential customers ([Jander, 2024](#)).

WTP-based approaches also help combat the phenomenon of “buy-down”, where passengers opt for cheaper fare classes even when they would have been willing to pay more for added value. By understanding WTP, airlines can personalise bundles and ancillary offers, increasing both conversion rates and customer satisfaction ([Shaw et al., 2021](#)).

Furthermore, WTP modelling supports the implementation of dynamic pricing strategies, allowing airlines to respond in real time to demand fluctuations, competitive actions, and contextual factors ([JetMate, 2025](#)).

2.2.3 Determinants of Willingness to Pay

Passenger WTP in aviation is shaped by a complex interaction of demographic, psychological, and contextual factors. Demographics such as age, income, and travel frequency play a significant role, with younger, higher-income, and frequent travellers often exhibiting a greater willingness to pay for premium services and sustainability initiatives ([Cordes et al., 2024](#)).

Service attributes, including comfort, convenience, and perceived value, are also important drivers. For example, passengers may be willing to pay more for extra legroom, priority boarding, or flexible ticketing options ([Chiambaretto, 2021](#)).

Psychological factors, such as attitudes toward risk and environmental consciousness, further influence WTP, particularly in the context of green premiums and voluntary carbon offsets ([Xu et al., 2022](#)).

Contextual elements, including the purpose of the trip, route length, and broader market conditions, can also affect WTP, sometimes dramatically, as seen during the COVID-19 pandemic ([Buchmann, 2025](#)).

2.2.4 Key Studies and Findings

Recent studies on willingness to pay (WTP) in aviation combine distinct methodological approaches, but collectively reveal how passenger choices are shaped by a mix of service experience, demographic, and psychological factors. [Warnock-Smith et al. \(2017\)](#) investigated how elements such as flight length, airline reputation, and service convenience were linked to higher WTP for ancillary services, showing how premium offerings and longer sectors drive improved passenger value. Building on the behavioural dimension, [Zhou et al. \(2020\)](#) focused on Chinese air travellers' seat selection, revealing a relation-

ship between price sensitivity and WTP and finding that ticket price perceptions, airline reputation, and convenience were all key drivers. Taking a different approach, [Xu et al. \(2022\)](#) explored not only the economic aspects but also the environmental and psychological determinants behind WTP for sustainable aviation fuel, highlighting the influence of factors such as environmental attitude, perceived efficacy, and transparency.

Author and Paper	Methodologies	Findings
Warnock-Smith et al. (2017) <i>“An analysis of ongoing trends in airline ancillary revenues”</i>	<ul style="list-style-type: none"> - Empirical analysis - Cross-sectional data review - Regression modelling 	<ul style="list-style-type: none"> - Longer flights and reputable carriers led to up to 30% higher WTP for ancillaries - Convenience and digital access (e.g. on-board Wi-Fi) significantly increased uptake - Ancillary revenue per passenger rose 268% over 10 years
Zhou et al. (2020) <i>“Willingness to pay for economy class seat selection: From a Chinese air consumer perspective”</i>	<ul style="list-style-type: none"> - Survey of 710 Chinese consumers - Discrete choice modelling 	<ul style="list-style-type: none"> - Price sensitivity negatively affected WTP for seat selection - Airline reputation and perceived flight convenience increased WTP - WTP varied by booking channel and timing, averaging CNY 80 ($\approx$ USD 12) for seat selection
Xu et al. (2022) <i>“Sustainable commercial aviation: What determines air travellers’ willingness to pay more for sustainable aviation fuel?”</i>	<ul style="list-style-type: none"> - Online survey (N = 1008, UK passengers) - Ordered logistic regression 	<ul style="list-style-type: none"> - WTP for sustainable aviation fuel driven by environmental values, risk perception, social trust, and transparency - Younger and more environmentally conscious groups reported a WTP premium averaging £25–40 per ticket

Table 2.2. Key studies and findings on WTP in aviation.

2.2.5 Challenges and Criticisms

Despite its promise, the use of WTP in aviation faces several challenges:

- **Measurement Bias:** Survey-based methods may overstate or underestimate true WTP due to hypothetical bias, social popularity, or lack of real economic consequences ([Hamdani et al., 2021](#)).
- **Data Limitations:** Revealed preference approaches require rich datasets and may not capture latent demand for unoffered services ([Jander, 2024](#)).
- **Dynamic Preferences:** Passenger WTP is not static; it fluctuates with macroeconomic conditions, competitive actions, and even global events such as pandemics ([Buchmann, 2025](#)).
- **Customer Perception:** Aggressive WTP-based pricing may alienate passengers if perceived as unfair or opaque, potentially harming brand loyalty ([Shaw et al., 2021](#)).

2.2.6 Future Directions

The future of WTP research and application in aviation is shaped by several trends:

- **AI-Driven Personalisation:** The adoption of advanced machine learning and AI tools will enable airlines to estimate WTP with greater precision and adapt offers in real time, moving closer to true offer optimisation and dynamic retailing ([Liu et al., 2024](#)).
- **Sustainability and Social Responsibility:** As environmental concerns grow, airlines will increasingly use WTP insights to design and price green products, such as SAF surcharges and carbon offsets, targeting segments most willing to pay for sustainability ([Hui and O'Connell, 2024](#)).

- **Integration with Ancillary Revenue Strategies:** WTP will become even more central to the unbundling and personalisation of ancillary services, allowing airlines to maximise revenue from each passenger segment ([Shaw et al., 2021](#)).
- **Improved Data Collection:** The proliferation of digital touchpoints and direct distribution channels will provide richer data for WTP estimation, supporting more granular segmentation and pricing ([JetMate, 2025](#)).

2.3 Survey-Based WTP Measurement

Survey-based approaches to determining WTP are fundamental to aviation research, providing airlines and academics with information on how much passengers are prepared to pay for various services and features. As airlines pursue more personalised pricing and product strategies, robust survey methodologies play an important role in capturing complex preferences and attitudes.

This subsection examines the design, implementation, and interpretation of survey-based WTP studies in aviation, emphasising the use of Likert-scale instruments and stated preference techniques.

2.3.1 Survey Design and Methodology

Survey-based WTP measurement in aviation generally uses stated preference (SP) and contingent valuation (CV) approaches. SP studies give respondents hypothetical scenarios, involving different combinations of service attributes and prices, asking them to make choices or indicate preferences ([Lindhjem and Navrud, 2012](#)). This method is especially useful for assessing new or hypothetical products, such as innovative ancillaries or sustainability features.

By contrast, CV directly asks participants about the maximum amount they would pay for a specific service, using open-ended questions, payment scales, or bidding games ([Lindhjem and Navrud, 2019](#)). Both methods are designed to obtain the upper limit of a respondent's WTP, resulting in useful information for pricing and product development.

Likert scales are commonly incorporated into these surveys to measure attitudes, perceptions, and behavioural intentions related to WTP. Respondents might rate their agreement with statements like “I am willing to pay extra for priority boarding” on a five- or seven-point scale ([Hamdani et al., 2021](#)). This allows researchers to capture both the monetary and psychological dimensions of WTP.

2.3.2 Implementation in Aviation Contexts

In aviation, WTP surveys are typically distributed at airports, online, or via airline customer databases. Sampling strategies are important for ensuring that findings are representative of the whole passenger base. Researchers look for diversity in demographics, travel frequency, and route types to improve the generalisability of results ([Hui and O’Connell, 2024](#)).

Survey scenarios are constructed to reflect realistic travel decisions. For example, passengers may be asked to consider a flight with specific features (such as extra baggage, seat selection, or eco-friendly options) and indicate their WTP for each. Conjoint analysis is often used to estimate the marginal WTP for individual attributes by analysing choices across multiple hypothetical offerings ([Hinnen et al., 2015](#)).

Surveys also frequently include questions on perceived value, satisfaction, and loyalty, allowing researchers to link WTP to broader constructs and support more effective segmentation and targeting by airlines ([Hui and O’Connell, 2024](#)).

2.3.3 Data Analysis and Interpretation

Analysis of survey-based WTP data involves descriptive statistics to summarise responses and inferential techniques such as regression to explore relationships between WTP and explanatory variables like income, travel frequency, or environmental attitudes. Advanced models, including hierarchical Bayes estimation and mixed logit models, help account for preference heterogeneity and estimate individual-level WTP ([Hinnen et al., 2015](#)).

Researchers must also address potential biases, such as hypothetical bias (where stated WTP differs from actual behaviour) and strategic bias (where respondents may misrepresent their true WTP). Careful survey design, validation questions, and, where possible, calibration with revealed preference data are used to mitigate these issues (Rocheva et al., 2022).

2.3.4 Key Studies and Findings

Research on WTP in aviation combines diverse methodologies and regional contexts, together advancing our understanding of how passenger valuations are shaped by psychological, environmental, and market-driven factors. Molin et al. (2017) examined how safety perceptions and attitudes influence WTP, highlighting the significant correlation between risk sensitivity and passenger payment preferences. Hinnen et al. (2015) shifted focus to the environmental dimension, employing adaptive conjoint analysis to reveal specific groups of Swiss air travellers who are willing to pay a premium for green products and carbon offsets—measured consistently with Likert-scale instruments. Lestari et al. (2019) introduced a mixed-methods framework, blending interviews with quantitative surveys to better estimate WTP for air travel in Indonesia, demonstrating how combining methodological approaches leads to richer insights into consumer decision-making. Hamdani et al. (2021) added a contemporary dimension by exploring WTP for low-cost flights during the COVID-19 pandemic, identifying perceived price fairness and psychological comfort as decisive elements.

Author and Paper	Methodologies	Findings
Molin et al. (2017) “Willingness to pay for safety improvements in passenger air travel”	<ul style="list-style-type: none"> - Survey (Likert scale) - Regression modelling - Analysis of risk perception 	<ul style="list-style-type: none"> - Risk-averse passengers had WTP up to 15% higher for added safety features - Fear of flying and higher income drove greater WTP - Likert scores for “safety concern” significantly correlated with higher WTP

Hinnen et al. (2015) “Willingness to pay for green products in air travel: Ready for take-off?”	<ul style="list-style-type: none"> - Adaptive choice-based conjoint survey - Five-point Likert scale - Swiss air travellers 	<ul style="list-style-type: none"> - 48% of respondents willing to pay a premium for carbon offsets - Mean WTP for green products increased by 15-20€ per ticket - Identified segments with strong eco-preferences using Likert item clustering
Lestari et al. (2019) “Analysis of ability to pay and willingness to pay of aircraft passengers toward Jakarta Surabaya high-speed train”	<ul style="list-style-type: none"> - Mixed-method questionnaire and interviews - WTP estimation - Indonesian passenger sample 	<ul style="list-style-type: none"> - Survey + interview results provided more accurate and higher WTP estimates vs. single method - WTP was $\approx 7-13\%$ above current average ticket pricing - Education and income were prominent WTP drivers
Hamdani et al. (2021) “Are passengers willing to pay extra money for domestic flight low-cost carrier airline in covid-19 pandemic?”	<ul style="list-style-type: none"> - Online survey - Likert scale for attitude and price perception - Statistical correlation analysis 	<ul style="list-style-type: none"> - Perceived price and psychological comfort were strongest WTP predictors during pandemic - WTP for extra safety measures rose by up to 19% in LCC markets - Attitude scores predicted additional payment propensity

Table 2.3. Key studies and findings on survey-based WTP measurement.

2.4 ML for Consumer Behaviour Prediction

Machine learning (ML) has become a transformative tool for predicting and understanding consumer behaviour in aviation. As airlines face greater competition and changing passenger expectations, ML facilitates the extraction of relevant data from large and complex datasets. These advanced analytical approaches allow airlines to move beyond static segmentation and traditional statistical models, supporting real-time and data-driven decisions that better reflect dynamic passenger preferences and market trends ([Lin, 2025](#)).

2.4.1 Machine Learning Approaches in Aviation

A wide range of ML techniques are applied to consumer behaviour prediction in aviation. Supervised learning algorithms (including decision trees, random forests, support vector machines, and neural networks) are employed to classify passengers, forecast demand, and predict ancillary purchases or upgrades. These models are trained on historical booking and transactional data, allowing airlines to identify patterns and drivers of consumer decisions ([Puranik et al., 2020](#)).

Unsupervised learning methods, such as k-means and hierarchical clustering, are used in market segmentation to identify latent passenger groups with similar behaviours or preferences, allowing for more targeted marketing and product bundling. Reinforcement learning, meanwhile, is increasingly used for dynamic pricing and offer optimisation, enabling airlines to adjust prices and offers in real time to maximise revenue ([Jasra et al., 2025](#)).

Recent studies also highlight the growing use of deep learning architectures, such as recurrent neural networks and multilayer perceptrons, which are particularly effective in modelling sequential and high-dimensional consumer data. These models often outperform traditional approaches, especially when large and complex datasets are available ([Maheshwari et al., 2018](#)).

2.4.2 Data Sources and Feature Engineering

ML models in aviation draw on diverse data sources, including booking histories, loyalty program activity, website and app interactions, and even unstructured data such as customer reviews and social media posts. Effective feature engineering is crucial, as it improves the predictive power of models (El-Hage, 2020). Features commonly used include booking lead time, travel frequency, ancillary purchase history, route characteristics, and contextual variables like seasonality and competitive actions (Workie et al., 2024).

Natural language processing is increasingly applied to analyse unstructured text data, such as customer feedback and online reviews. This provides additional sentiment and purpose signals, which further refines consumer behaviour predictions (Lin, 2025).

2.4.3 Model Evaluation and Interpretability

Robust evaluation is essential for ML models. Common metrics include accuracy, precision, recall, F1 score, and area under the ROC curve (AUC), depending on the prediction task (Lin, 2025). Cross-validation and out-of-sample testing are also standard practices to ensure model generalisability and to prevent overfitting (El-Hage, 2020).

Interpretability is also a key concern, especially when ML outputs inform pricing or customer targeting. Techniques such as feature importance rankings and SHAP values are used to provide transparency, helping stakeholders understand the factors driving model predictions (Nanyonga et al., 2024).

2.4.4 Key Studies and Findings

Recent literature on the application of ML to WTP and consumer behaviour in aviation highlights the diversity of techniques and their practical effectiveness in real-world airline settings. El-Hage (2020) explored the classification of seat reservation purchase behaviour, showing how standard neural and support vector machine models can categorise consumer intent based on historical data. AlHabbal (2022) introduced a comparative

approach, using algorithms from K-Nearest Neighbour and Random Forest to advanced deep learning, achieving high predictive accuracy in satisfaction and churn, indicating ML's value in optimising passenger outreach. [Shahid et al. \(2024\)](#) extended these concepts to loyalty prediction, integrating sentiment analysis to reveal how emotional and experiential data can boost forecast accuracy for return visits. Finally, [Workie et al. \(2024\)](#) turned to commercial operations, using a suite of regression and ensemble models to improve the precision of in-flight sales predictions, with XGBoost consistently outperforming traditional approaches.

Author and Paper	Methodologies	Findings
El-Hage (2020) <i>“Predicting purchase of airline seating using machine learning”</i>	<ul style="list-style-type: none"> - MLP & SVM classifiers - Data from seat reservation logs - Feature selection & numerical encoding 	<ul style="list-style-type: none"> - ML models classified seat booking intent with $\approx 70\%$ accuracy - Lead time, price, and prior behaviour were key features - MLP slightly outperformed SVMs in most scenarios
AlHabbal (2022) <i>“Predicting & optimising airlines customer satisfaction using classification”</i>	<ul style="list-style-type: none"> - KNN, RF, Decision Trees, CNN-LSTM - Classification on customer satisfaction & churn - Comparative analysis 	<ul style="list-style-type: none"> - Deep learning (CNN-LSTM) achieved up to 94% classification accuracy for satisfaction/churn - RF consistently performed well for tabular data - Model interpretability improved by feature importance analysis

Shahid et al. (2024) <i>“Predicting customer loyalty in the airline industry: A machine learning approach integrating sentiment analysis and user experience”</i>	<ul style="list-style-type: none"> - XGBoost & sentiment analysis - Survey & feedback integration - Loyalty (repeat purchase) forecasting 	<ul style="list-style-type: none"> - Model achieved up to 85% accuracy in loyalty prediction - Sentiment scores from reviews boosted forecast precision - User experience metrics were strong loyalty predictors
Workie et al. (2024) <i>“In-flight sales prediction using machine learning”</i>	<ul style="list-style-type: none"> - Linear Regression, Decision Tree, RF, XGBoost - Actual in-flight sales datasets - Feature importance mapping 	<ul style="list-style-type: none"> - XGBoost produced highest sales prediction accuracy (MAE up to 15% lower than baseline) - Product popularity and flight segment were top predictors - ML-based predictions allow dynamic sales and inventory planning

Table 2.4. Key studies and findings on ML for consumer behaviour.

2.5 Gap Analysis

Despite the fast evolution of ancillary services as an important source of income for airlines and the increasing use of ML in consumer behaviour prediction, several important gaps persist in the literature, especially at the intersection of route-specific analysis, advanced analytics, and survey-based willingness to pay measurement. Existing research has widely documented the growth and strategic relevance of ancillary revenues, as well as the major determinants of WTP for such services, often focusing on aggregate trends across airlines, regions, or service types. However, most studies rely on generalised datasets or revealed preference data, paying little attention to the complex, route-level dynamics that impact passenger choices in specific market circumstances. This is especially true for short-

haul routes like Barcelona–London Gatwick (BCN-LGW), where competitive pressures, passenger profiles, and ancillary preferences may differ significantly from long-haul or intercontinental markets.

Moreover, while survey-based methods (particularly those employing Likert-scale instruments) are recognised as robust tools for capturing attitudinal and psychological drivers of WTP, their integration with advanced ML models remains underexplored in aviation research. The majority of prior work either uses survey data for descriptive or econometric analysis, or applies ML to large-scale transactional data, but rarely combining the strengths of both approaches. This creates a methodological gap: there is little empirical evidence on how numerically encoded Likert-scale responses, when collected and postprocessed, can be used by ML algorithms such as XGBoost, Random Forest, or Support Vector Machines to accurately predict WTP for ancillary services at the individual passenger level.

Additionally, studies that do apply ML to passenger data often lack interpretability, offering little information into which demographic or behavioural features most influence WTP predictions. The use of explainability techniques (such as SHAP (SHapley Additive exPlanations) or permutation importance) has only recently gained traction, and their application to survey-based WTP prediction in aviation is still emerging. This limits the practical utility of ML models for airlines, who need clear and concrete data to tailor supplementary products and marketing tactics to specific passenger segments.

Finally, there is a lack of concrete studies focused on the ancillary services on short-haul routes. While some exploratory works have considered third-party ancillary revenues and passenger willingness to pay, these studies typically aggregate findings across both short- and long-haul segments, rarely isolating the unique characteristics and challenges of short-haul markets and making difficult to draw useful conclusions about ancillary service preferences, pricing sensitivity, or passenger segmentation specific to these routes. As a result, there remains a significant need for research that directly addresses ancillary service strategies, passenger behaviour, and WTP determinants within the short-haul context.

In summary, this thesis addresses a critical gap by integrating Likert-scale survey data with advanced ML techniques to predict passenger WTP for ancillary services on the BCN-LGW route. It also directly tackles the absence of concrete studies focused on ancillary service dynamics in the short-haul segment, providing an interpretable and empirical framework that not only forecasts individual WTP but also segments the passenger base for targeted ancillary offerings.

Chapter 3

Methodology

3.1 Research Design

Every research project requires a clear plan, and the ‘research onion’ model proposed by [Sanders \(2023\)](#) provides exactly that. This section breaks down its application: the chosen strategy, the methodological choices employed, the defined time horizon, and the main techniques and procedures used to achieve the research objectives.

Research Strategy

This study uses a **survey strategy**, meaning that data is collected directly from passengers through a structured questionnaire. The survey helps gather consistent information about passengers’ demographics, travel habits, and their willingness to pay for extra services on the BCN-LGW route.

Other research strategies, such as experiments, case studies, and ethnography, were considered but not selected. Experimental designs were discarded because manipulating variables like pricing or service availability is impractical and raises ethical concerns in a real-world airline environment. Case studies were not chosen as they would focus too narrowly on a single airline or route, reducing the generalisability of findings. Ethnographic approaches were also rejected since they rely on observation and qualitative data, which do not provide the standardised, quantitative information needed for predictive modelling with machine learning.

Research Choices

A **mono-method quantitative approach** is used in this research. This means that only numerical data from the survey is collected and analysed, without any qualitative interviews or open-ended questions. This choice makes it easier to use statistical and machine learning tools to analyse the data.

Mixed-methods and qualitative-only approaches were considered but ultimately set aside. Incorporating qualitative interviews or open-ended survey questions could have provided deeper insights into passenger motivations, but these methods were excluded to keep the study focused on numerical data suitable for statistical and machine learning analysis.

Time Horizons

The research has a **cross-sectional time horizon**. All data is collected at one specific point in time, giving a snapshot of how passengers feel about paying for ancillary services during the study period, rather than tracking changes over time.

Longitudinal and diary-based time horizons were evaluated but not chosen for this research. While tracking passengers over multiple flights or collecting repeated measures could reveal how willingness to pay changes over time, these approaches were deemed too resource-intensive and logistically challenging, especially given the need to maintain participant anonymity and the likelihood of participant dropout. The cross-sectional approach was therefore preferred for its efficiency and practicality.

Techniques and Procedures

The study starts by designing and distributing a Likert-scale survey to passengers. The responses are turned into numbers and combined to create scores for each extra service. The data is then cleaned and prepared for analysis. Several machine learning models—such as XGBoost, Random Forest, Logistic Regression, and SVM—are used to predict willingness to pay. Feature importance methods like SHAP and permutation importance help explain which factors matter most. Finally, k-means clustering is used to group passengers with similar payment preferences, and the models are tested for accuracy and reliability using standard validation techniques.

Several alternative techniques and procedures were considered but not adopted. Manual statistical analysis, such as using only basic regression models, was rejected due to its limited predictive power and inability to handle complex, non-linear relationships in the data. Qualitative coding and thematic analysis of open-ended responses were also discarded, as they do not produce the structured data required for machine learning. Additionally, deep learning models like neural networks were considered but ultimately excluded due to their need for larger datasets, greater computational resources, and lower interpretability compared to algorithms like XGBoost and Random Forest, which are more suitable for this study's objectives.

3.2 Survey Design

The survey instrument for this study was developed to capture the key demographic, behavioural, and attitudinal variables that influence passengers' willingness to pay for ancillary services on the short-haul BCN-LGW route.

The objective was to study a short-haul route, and the BCN-LGW connection was selected because it is one of Europe's most popular and well-trafficked city pairs ([OAG, 2025](#)). Additionally, given my personal network as a Barcelona native currently studying in the UK, this route is one that most of my contacts have flown, which facilitates effective survey distribution.

The design is based on established academic literature and best practices in air transport and consumer behaviour research, ensuring both validity and comparability with prior studies. The full survey questionnaire is provided in [Appendix B](#).

Demographic and Travel Profile Variables

The following variables were included in Section A to define the passenger profile:

- **Gender**
 - Choices: Male, Female, Prefer not to say
 - Justification: Gender is a standard demographic variable in air transport and consumer behaviour research, consistently used in both academic and industry

surveys to capture potential differences in preferences and WTP. The inclusion of “Prefer not to say” ensures inclusivity and aligns with ethical research standards recommended by the UK CAA and government surveys. ([UK CAA, 2024](#); [UK Gov., 2018](#); [Zhou et al., 2022](#))

- **Age**

- Choices: 18–24, 25–34, 35–44, 45–54, 55–64, 65+
- Justification: These age bands are widely used in aviation and tourism studies, reflecting generational and life-stage differences in travel behaviour and WTP. The bands match those recommended by UK CAA, Eurostat, and OAG for comparability and to capture generational shifts (e.g., Gen Z, Millennials, Gen X, Boomers) ([OAG, 2024](#); [Urhausen, 2008](#)). Respondents under 18 are excluded as only adults are surveyed.

- **Annual Income (€)**

- Choices: < 20,000, 20,000–34,999, 35,000–49,999, 50,000–74,999, 75,000–99,999, ≥ 100,000
- Justification: Income is a key predictor of WTP for airline ancillaries. The brackets are based on European income distributions, using a semi-logarithmic scale to capture the typically skewed nature of income and to try to ensure enough respondents per group ([INE, 2023](#); [UK Gov., 2022](#)). Prior research confirms higher WTP among higher-income groups ([Cordes et al., 2024](#)).

- **Travel Purpose**

- Choices: Leisure, Business, Visiting family/friends, Education
- Justification: Journey purpose is a critical determinant of ancillary WTP, with business travellers typically more willing to pay for comfort and flexibility. These categories are standard in airline and tourism research. ([Kassir and Ashaal, 2021](#); [Liu et al., 2024](#))

- **Travel Frequency**

- Choices: 0–1, 2–4, 5–7, ≥ 8 flights per year
- Justification: Frequency of air travel is a key segmentation variable, as frequent flyers have different ancillary needs and WTP. These brackets are aligned with industry practice and allow for meaningful segmentation ([Cordes et al., 2024](#); [Liu et al., 2024](#)).

Ancillary Service Selection

For Section B, the selection of ancillary services is based on their relevance and revenue-generating potential for short-haul routes such as BCN-LGW. The four included are:

- Seat Selection
- Checked Baggage
- In-Flight Meals
- Wi-Fi

This selection is supported by recent academic and industry studies, which identify seat selection, baggage, and meals as among the most valued and purchased ancillaries on short-haul flights, with Wi-Fi rapidly growing in importance as airlines install connectivity across their fleets ([Škurla Babić et al., 2019](#)). A 2023 Viasat survey found that 81% of airline passengers said Wi-Fi was important to their onboard experience, and that Wi-Fi is now considered more desirable than meals or in-flight entertainment on short-haul flights ([Viasat, 2023](#)). Also, industry reviews consistently report that these four ancillaries are both widely offered and highly relevant for passenger experience and airline revenue management on intra-European routes ([Ozmec-Ban et al., 2022](#)).

Survey Structure

- **Section A:** Captures demographic and travel profile information using closed-ended questions with pre-defined categories, justified above.

- **Section B:** Assesses attitudes and trade-offs regarding each ancillary service using 5-point Likert-scale statements (1 = Strongly Disagree, 5 = Strongly Agree) ([Azad et al., 2023](#); [Zhou et al., 2022](#)). For each ancillary, two attitudinal and one trade-off questions are included, allowing for the computation of a composite WTP score per service when adding the results.

3.3 Data Collection

Data for this study was collected using an online survey administered via Google Forms. To maximise accessibility and response rates among potential participants, two versions of the survey were distributed: one in English and another in Spanish. This bilingual approach ensured that both English-speaking and Spanish-speaking respondents could participate comfortably.

The survey was open for a period of four weeks, during which time responses were gathered using a snowball sampling technique ([Sanders, 2023](#)). The initial distribution targeted my personal and academic networks, with participants being encouraged to share the survey further within their own circles to broaden the reach. After the survey period, responses from both language versions were merged into a single dataset for analysis, with all responses anonymised to maintain participant confidentiality. This approach facilitated efficient data collection and provided a robust sample for subsequent quantitative analysis.

The expected number of survey responses was approximately 200; in practice, a total of 179 responses were received. Although the final number was slightly lower than anticipated, it remains sufficiently robust for the intended quantitative and machine learning analyses, allowing for reliable model training.

3.4 Data Preprocessing

Data preprocessing is a critical stage in any data-driven research, as the quality and structure of the dataset directly impact the reliability of subsequent analysis and model predictions (Kotsiantis et al., 2006).

Initially, responses were reviewed for class representation within the "travel purpose" variable, as it existed a significant imbalance in the distribution of responses. The categories "Education" and "Visiting family/friends" were heavily underrepresented, accounting for only 3 and 5 responses respectively (less than 5% of the total dataset). Statistical best practices recommend that categories with very low sample sizes should be excluded from analysis, as their inclusion can introduce significant bias and reduce the validity of any following statistical or machine learning models (Hair, 2014). Consequently, these two categories were removed from the final dataset, keeping only the "Leisure" and "Business" traveller groups, which offered an adequate sample size for further analysis.

To prepare the data for quantitative analysis and input into predictive models, all Likert-scale responses were numerically encoded in accordance with standard conventions (1 = Strongly Disagree to 5 = Strongly Agree). This not only facilitated the computation of descriptive statistics but was also needed to meet the input requirements of most machine learning algorithms, which operate on numerically represented features (Field, 2024).

So, to quantify willingness to pay for each ancillary service, a composite WTP score was constructed. Recognising that the survey items varied in their directness regarding monetary trade-offs, additional weight was assigned to the item that most explicitly addressed willingness to pay (specifically, questions SS3, CB3, IM2, and WF2). The formula developed for this composite score is as follows:

$$WTP_{ancillary} = \frac{Q_1 + Q_2 + 2 \cdot Q_3}{20} \cdot 5$$

where Q_1 and Q_2 represent subjective or perception-oriented responses, and Q_3 represents the response explicitly concerned with economic trade-off. Division by 20 ensures normalisation (as the maximum achievable score is 20), while multiplying by 5 rescales the index, keeping consistency with the original Likert-scale and ensuring interpretability.

Finally, to enable effective ML modelling and interpretation, the original composite WTP scores were transformed into three distinct categories: Low, Medium, and High WTP. Specifically, responses were categorised as follows: scores of 1–2.5 were labelled as “Low”, a score of 2.5–3.5 as “Medium”, and scores of 3.5–5 as “High” willingness to pay.

This categorisation offers several advantages: it simplifies the prediction problem to a multiclass classification task, improves interpretability for statistical analysis, and aligns with common practices in social sciences, where Likert-scale data are often discretised to facilitate reliable group comparisons (Altman and Royston, 2006; Norman, 2010). This approach also addresses challenges related to the ordinal (but not interval) nature of Likert responses, supporting more robust model evaluation using widely accepted classification metrics.

3.5 Machine Learning Models

Once the data has been preprocessed, the next step is selecting the right machine learning models, as they play an important role in understanding the dynamics behind passengers’ WTP. By using different algorithms, this thesis can explore the data from multiple perspectives, achieving a balance between predictive power and interpretability.

3.5.1 Justification for Model Choices

As mentioned, the ML models chosen for this study were selected to provide both predictive accuracy and interpretability when analysing passengers’ WTP for ancillary services using survey data.

Logistic Regression is a well-established method in social science and survey research. Its strength lies in its transparency and the ability to directly interpret feature coefficients, which is particularly important for understanding how specific demographic and behavioural variables influence binary outcomes like purchase intent (Zaidi and Al Luhayb, 2023).

Random Forest was included because of its capacity to handle complex, non-linear relationships within the data, its robustness to overfitting, and its demonstrated success in identifying influential predictors in social surveys (Dureh et al., 2018). This algorithm’s built-in mechanism for evaluating variable importance provides useful information about which factors have the greatest influence on WTP across varied respondent groups.

Support Vector Machines (SVM) improves the model suite, particularly in scenarios where the number of features is large in relation to sample size. SVMs are especially useful for their capability to construct optimal hyperplanes in high-dimensional spaces, catching subtle class boundaries that simpler models might miss. Their reliability under the “curse of dimensionality” makes them a strong fit for survey datasets containing numerous variables drawn from diverse respondent profiles (Pisner and Schnyer, 2020).

Finally, **XGBoost** was selected because of its superior predictive power and its resilience to missing values and mixed data types often found in real-world survey datasets (Zhang et al., 2022). Its gradient boosting framework not only improves accuracy over base learners but also natively handles missing data during training, reducing preprocessing complexity and supporting robust outcomes.

3.5.2 Theoretical Background

A clear understanding of the theoretical basis of each ML model is necessary for interpreting results and assessing the suitability of each approach within our framework. Below, each algorithm is briefly introduced with its core principles and ideas.

Logistic Regression

Logistic Regression is designed for binary classification, modelling the probability that an outcome belongs to a given category. It estimates the log-odds of the target as a linear function of input variables:

$$\text{logit}(p) = \log \left(\frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n \quad (3.1)$$

where p is the predicted probability and x_i are the features. The model uses the logistic (sigmoid) function to constrain the output between 0 and 1:

$$p = \frac{1}{1 + e^{-(\beta_0 + \sum \beta_i x_i)}} \quad (3.2)$$

The coefficients β are estimated through maximum likelihood estimation, offering direct interpretability of feature effects in terms of odds ratios (Lee, 2020). Logistic regression is widely used for survey research when interpretability and inference are key.

Random Forest

Random Forest is an ensemble learning method that generates a large collection of decision trees, each trained on a random subset of the data and features, and the final classification is determined by majority vote across trees. At each node of a tree, the algorithm chooses the split that maximises information gain (often using Gini impurity or entropy):

$$Gini = 1 - \sum_{i=1}^C p_i^2 \quad (3.3)$$

where p_i is the probability of an item being classified into a class i . Randomness in sampling and feature selection makes individual trees less correlated, reducing variance and minimising overfitting (Biau and Scornet, 2016). This method handles complex, non-linear interactions and naturally estimates variable importance.

Support Vector Machine (SVM)

SVM is a powerful classification method seeking the hyperplane that best separates two classes with the maximum margin. For linearly separable data:

$$\text{Find } \mathbf{w}, b \text{ to maximise margin subject to } y_i(\mathbf{w} \cdot x_i + b) \geq 1 \quad (3.4)$$

Here, y_i is the label, x_i is the feature vector, and $\mathbf{w}$ and b define the hyperplane. When separation is not possible in the original space, a kernel function (e.g., RBF, polynomial) projects data into higher dimensions, and the optimal solution is found using convex

optimisation. SVMs are robust in high-dimensional spaces and resistant to overfitting, especially in settings with many features relative to observations ([Fletcher, 2008](#)).

XGBoost

XGBoost implements the gradient boosting framework, building trees sequentially so that each new tree corrects the errors of the combined ensemble so far. The prediction for an input x is:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F} \quad (3.5)$$

where K is the number of boosting rounds, and each f_k is a decision tree. The model minimises a regularised loss function incorporating both residual error and tree complexity:

$$L(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (3.6)$$

where l is a differentiable loss function and Ω is a regularisation term. XGBoost is highly efficient and effective with categorical, continuous, and missing values, and is known for its superior predictive ability ([GFG, 2025](#)).

3.6 Evaluation Metrics

Following the selection of machine learning models, the choice and interpretation of evaluation metrics become essential for reliably comparing and validating their performance. Since survey-based data often presents challenges such as class imbalance or varying error costs, accurately quantifying model performance requires a solid understanding of each metric's strengths and limitations.

3.6.1 Rationale for Metric Selection

A complete set of metrics offers a detailed view of model quality. **Accuracy** measures the overall rate of correct predictions and is calculated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.7)$$

where TP , TN , FP , and FN are true positives, true negatives, false positives, and false negatives, respectively. While intuitive, accuracy can be misleading in imbalanced datasets, as it may be high even if the model fails to identify the minority class (Naidu et al., 2023).

Precision reflects the proportion of positive identifications that were correct:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.8)$$

High precision is essential in contexts where false positives carry significant consequences (Naidu et al., 2023).

Recall (also known as sensitivity) quantifies the proportion of actual positives correctly identified:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.9)$$

A high recall indicates the model's ability to capture the real positive cases, minimising false negatives (Naidu et al., 2023).

F1 Score represents the harmonic mean of precision and recall:

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3.10)$$

By balancing these two metrics, the F1 score helps ensure that neither precision nor recall are excessively sacrificed. This is especially valuable in imbalanced datasets or when both types of misclassification errors (false positives and negatives) are relevant. For these reasons, this study prioritises the model with the highest F1 score as the "best", since it represents the ideal trade-off and robustness for real-world prediction scenarios (Naidu et al., 2023).

Finally, the **ROC-AUC** (Receiver Operating Characteristic – Area Under Curve) metric provides a threshold-independent assessment of model discrimination, plotting the true positive rate against the false positive rate at various thresholds. A ROC-AUC value close to 1.0 indicates excellent class separation, while a value near 0.5 suggests no discriminative ability (Naidu et al., 2023).

3.6.2 Confusion Matrix

After evaluating the performance metrics across our ML models and identifying the best-fitting model for each ancillary service with the F1 Score, the next step involves analysing the confusion matrix for the selected model in each case. A confusion matrix is a powerful tool for assessing classification performance, particularly in multi-class settings. For our study, which predicts passenger WTP categories (Low, Medium, or High) across different ancillary services, the confusion matrix offers a granular view of the model’s predictions versus actual classifications, revealing strengths and potential biases in its decision-making (Hoang, 2023).

The matrix is organised such that each row represents the instances in an actual class, while each column represents the instances in a predicted class. For a three-class classification problem, the matrix will be 3×3 , with the following structure:

	Predicted Low	Predicted Medium	Predicted High
Actual Low	...	...	...
Actual Medium	...	...	...
Actual High	...	...	...

Table 3.1. Example confusion matrix for a multi-class classification problem with Low, Medium, and High categories.

Each cell (i,j) in the matrix indicates the number of passengers whose true WTP category was class i (row) and who were predicted as class j (column). Diagonal elements represent correct classifications, while off-diagonal elements indicate misclassifications between categories.

This multi-class confusion matrix allows to identify specific patterns of misclassification (e.g., if Medium WTP passengers are often confused with Low or High), enabling targeted improvements to the model. Additionally, it provides the basis for calculating various performance metrics such as per-class precision, recall, and F1-score.

3.7 Feature Importance Analysis Methods

Once evaluated the ML models, the following step involves interpreting the results to extract meaningful information. Specifically, to deepen model interpretability, it is essential to identify which features most significantly influence predictions. Feature importance analysis serves this purpose by quantifying the contribution of different variables, thereby guiding both model refinement and business decision-making. Two widely adopted methods, SHAP (SHapley Additive exPlanations) values and Permutation Importance, offer complementary perspectives for assessing feature relevance in complex models, allowing for a more complete understanding of passenger WTP dynamics.

3.7.1 SHAP Values

On one hand, SHAP (Shapley Additive Explanations) values are a sophisticated approach grounded in cooperative game theory. SHAP assigns each feature an importance value for a particular prediction, reflecting how much that feature contributes to moving the prediction away from the average baseline. This method is model-agnostic and can explain a wide variety of algorithms, including tree-based, linear, and even deep learning models (Cooper, 2021).

The core advantage of SHAP lies in its ability to decompose a prediction into a sum of individual feature contributions. For any instance, the model's prediction can be written as:

$$f(x) = \text{base value} + \sum_{i=1}^n \text{SHAP value}_i \quad (3.11)$$

where the base value represents the average prediction if no features are known, and each

SHAP value quantifies a feature’s impact in the context of all possible feature combinations. SHAP sets itself apart by ensuring consistency and local accuracy, meaning the sum of feature attributions always equals the difference between the actual prediction and the baseline.

In practice, SHAP values help both in summarising global feature importance (by averaging the magnitude of SHAP values across the dataset) and in providing local insights (for instance, showing how the combination of input features in a specific instance drives its individual prediction). The interpretability extends to visualisations such as force plots and summary plots, which reveal not only the most influential features but also the direction (positive or negative) and intensity of their effects ([Molnar, 2019b](#)).

3.7.2 Permutation Importance

On the other hand, Permutation Importance is a robust, model-independent technique used to quantify the importance of each feature by measuring how much a model’s performance degrades when that feature’s values are randomly shuffled. By breaking the relationship between an individual feature and the target variable, one can directly observe the contribution of that feature to the model’s predictions ([Molnar, 2019a](#)).

The process begins with training a model and establishing a baseline performance metric, such as accuracy or mean squared error, on a validation or test set. Each feature is then permuted (shuffled) one at a time, and the change in performance is recorded. A large drop in performance indicates that the model heavily relies on the permuted feature; while a negligible change means limited importance. This procedure is repeated for every feature, resulting in a straightforward ranking of feature importances based on their impact on predictive accuracy ([Galli, 2024](#)).

Permutation Importance stands out for its versatility. It is entirely agnostic to model architecture and can be applied to any supervised learning model, including ensembles and black-box algorithms. Additionally, permutation-based analysis is especially useful in identifying redundant or irrelevant features and supports feature selection and data quality assessments ([Molnar, 2019a](#)).

However, Permutation Importance provides only a global measure of importance; it does not explain the contribution of features for individual predictions, nor does it clarify the directionality (i.e., whether a feature increases or decreases a specific prediction). Despite these limitations, its ease of implementation and conceptual clarity make it an essential part of ML interpretability toolkits ([Galli, 2024](#)).

3.8 Software and Tools

Finally, the development and implementation of this project used a combination of well-known software tools and programming resources, each selected for their reliability and suitability to specific tasks along the process.

For survey data collection, **Google Forms** was chosen due to its ease of use, cost-effectiveness, and easy integration with the Google ecosystem. Its intuitive interface allowed the quick design and distribution of customised surveys, while its automatic connection with Google Sheets made it simple to aggregate and analyse responses in real time. Also, Google Forms is accessible from any device and supports unlimited responses with a range of question types, ensuring efficient and secure collection of diverse participant data.

All programming was carried out using **Python**, selected for its simplicity, versatility, and extensive ecosystem of powerful libraries. Python is a leading language in both data science and machine learning, offering high readability and rapid prototyping. The coding environment was **Visual Studio Code (VS Code)**, chosen for its lightweight footprint, customisable interface, and excellent support for Python development through features like IntelliSense and integrated debugging. Also, VS Code's broad extension marketplace facilitated efficient programming, version control, and code navigation, significantly improving productivity.

Lastly, core ML model development and evaluation were handled with Python’s **scikit-learn** library, praised for its wide array of algorithms (classification, regression, clustering), comprehensive documentation, and user-friendly API. This allowed for fast, repeatable experiments and robust evaluation using the previously mentioned metrics. Data manipulation and preprocessing were performed using **Pandas** and **NumPy**, which deliver flexible data structures and high-performance numerical computation, essential for transforming survey data into actionable ML inputs.

The step-by-step workflow of the data analysis process is described in pseudocode in [Appendix C](#).

Chapter 4

Results

4.1 Descriptive Statistics

This subsection offers a detailed summary of the survey sample characteristics and the distributions of passengers' WTP for ancillary services on the BCN-LGW route. The aim is to clearly present the demographic and behavioural background of respondents and to summarise the key patterns in WTP responses using tables and targeted visualisations.

4.1.1 Sample Composition

The following analysis presents the demographic and travel behaviour profile of passengers who participated in the survey. Understanding the characteristics of the sample is necessary for evaluating the patterns in WTP for ancillary services in later sections. Table 4.1 contains the respondents' distribution across gender, age, annual income, usual travel purpose, and annual flight frequency. Note that percentages are based on the total number of final responses, which were 171.

Variable	Categories	Count	Percentage
Gender	Male	90	52.63%
	Female	81	47.37%
Age	18 - 24	38	22.22%
	25 - 34	36	21.05%
	35 - 44	31	18.13%
	45 - 54	25	14.62%
	55 - 64	22	12.87%
	65 +	19	11.11%
Annual Income	< €20,000	34	19.88%
	€20,000 - €34,999	42	24.56%
	€35,000 - €49,999	30	17.54%
	€50,000 - €74,999	26	15.20%
	€75,000 - €99,999	21	12.28%
	≥ €100,000	18	10.53%
Usual Travel Purpose	Leisure	103	60.23%
	Business	68	39.77%
Annual Flight Frequency	0 - 1 flights	49	28.65%
	2 - 4 flights	57	33.33%
	5 - 7 flights	38	22.22%
	≥ 8 flights	27	15.79%

Table 4.1. Sample distribution by demographic and travel behaviour variables.

As seen in Figure 4.1, the sample is nearly gender-balanced, with 52.6% male and 47.4% female respondents. This proportionality helps ensure that any findings on WTP are not disproportionately skewed by one gender's preferences. Balanced gender representation increases the reliability of gender-comparison analysis and mirrors the likely composition of many short-haul European flights, improving generalisability.

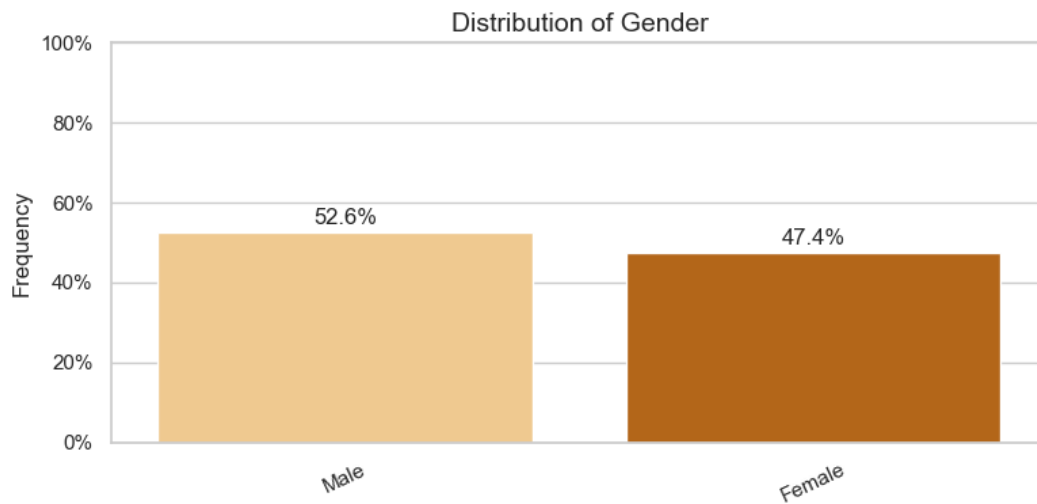


Figure 4.1. Distribution of respondents by gender.

Figure 4.2 demonstrates that respondents span all age brackets, forming a relatively smooth gradient from younger to older groups. The largest portions come from the 18–24 and 25–34 segments (a combined 43.3%), which is typical for short-haul, city-pair routes popular among students, young professionals, and early-career business travellers. Representation in senior categories is continuously decreasing, with only 11.1% aged 65 or older. Note that such a trend might reflect the potential survey accessibility differences.

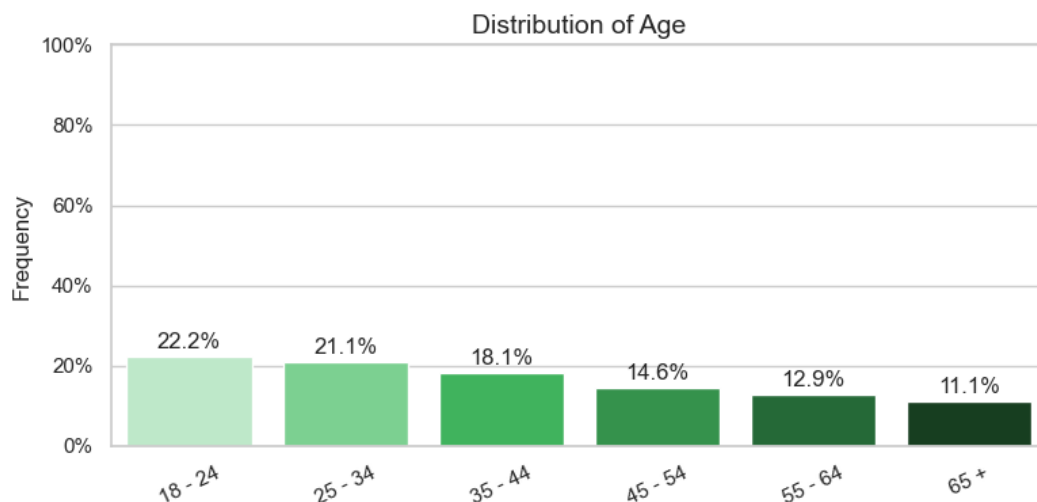


Figure 4.2. Distribution of respondents by age.

As illustrated in Figure 4.3, income distribution covers the full expected range but, as is common in social surveys, the highest brackets ($<€100,000$) have fewer respondents. The largest group, at 24.6%, earns €20,000–€34,999. Decreases in sample size with rising income are typical, as higher-income individuals are a smaller proportion of the general population and may be less likely to respond to surveys. However, the presence of a substantial number of high-income respondents (over 22% earning €75,000 or more) is expected to give enough analytical power to detect income effects on WTP for ancillary services.

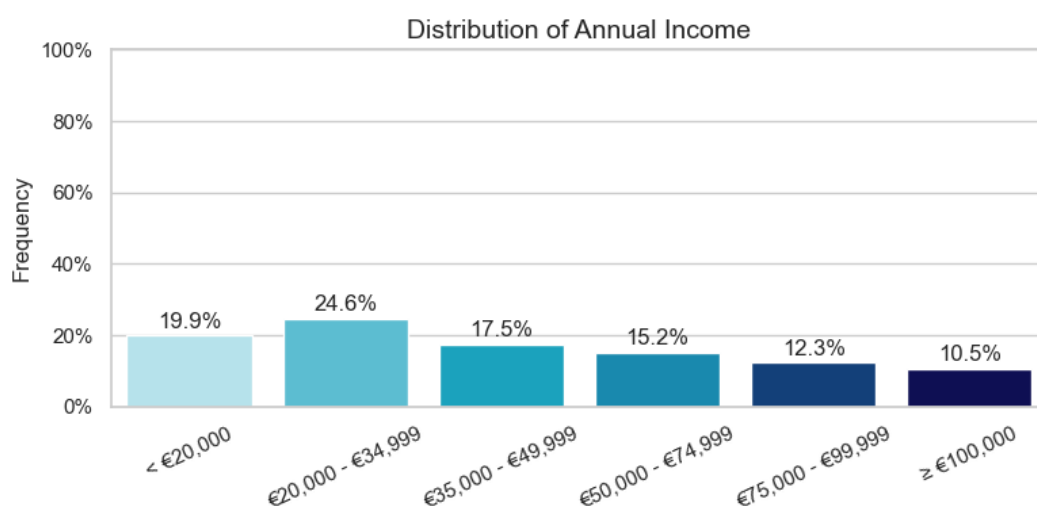


Figure 4.3. Distribution of respondents by annual income bracket.

According to Figure 4.4, the majority of passengers report leisure as their primary purpose (60.2%), while 39.8% indicate business, which is logical given that the survey was distributed primarily through personal networks rather than corporate channels. This diversity in travel purpose is important, as ancillary product valuations often differ between leisure and business segments; hence, comparative analysis by purpose is well supported.

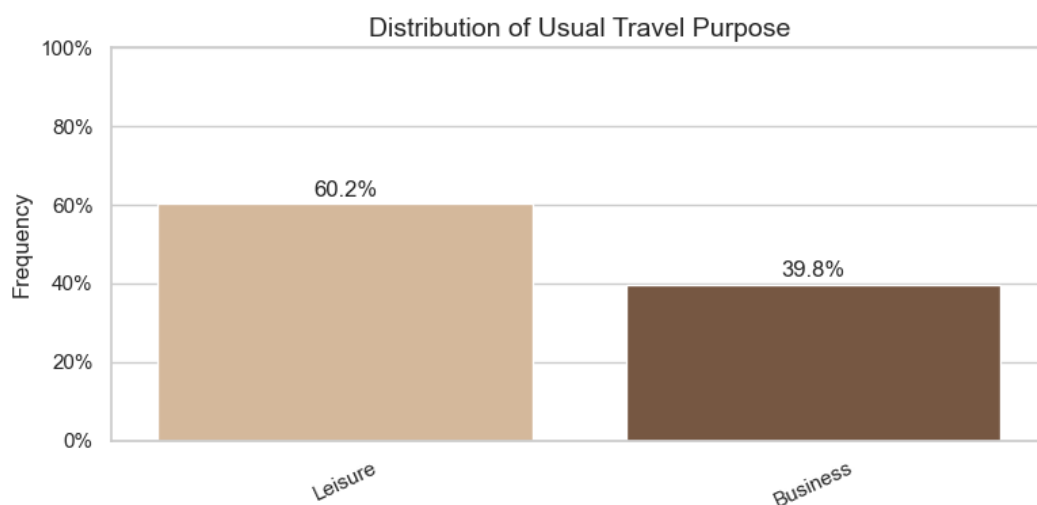


Figure 4.4. Distribution of respondents by usual travel purpose.

Figure 4.5 shows that the range of annual flight frequency is well captured: roughly one third are occasional flyers (0–1 trips/year), another third are “regulars” (2–4 times), with diminishing proportions as frequency increases. This is in line with expected air travel market segmentation, where frequent flyers form a smaller but significant customer base. Having distinct segments here allows for robust testing of whether ancillary valuations depend on how often a person flies.

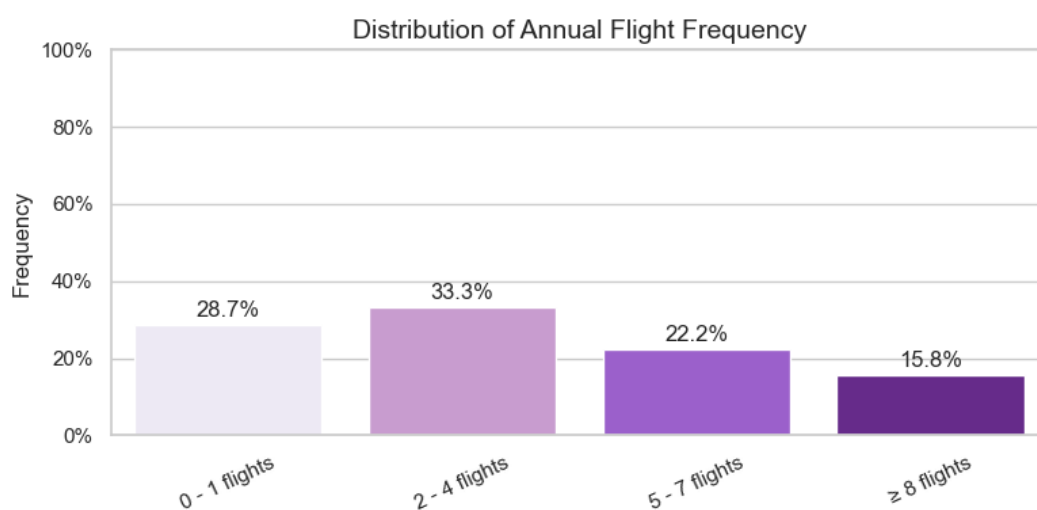


Figure 4.5. Distribution of respondents by annual flight frequency.

Overall, the study’s sample reveals a healthy diversity in core demographic and behavioural parameters. This ensures that the following analysis (particularly when predicting or segmenting WTP for ancillary services) rest on a representative and analytically rich foundation, as balanced distributed categories not only help avoid bias but also offer practical relevant information.

4.1.2 WTP for Ancillary Services

Passengers’ WTP for ancillary services on the BCN-LGW route varies significantly by product type. For each ancillary (seat selection, checked baggage, in-flight meals, and Wi-Fi), respondents’ composite WTP scores are illustrated using both histograms (to show overall distribution) and boxplots (to clarify central tendency and spread).

As illustrated in Figure 4.6, the histogram for seat selection reveals a generally positive skew, with many respondents reporting mid-to-high WTP scores. There is a prominent peak around scores of 3.5–4.0, indicating that an important portion of travellers are open to paying extra for their preferred seat. The boxplot supports these observations, with a median WTP score above 3 and an interquartile range of roughly 2.5 to 4.2. Also, while there are extreme low and high valuations, the majority of respondents fall somewhere in the middle, indicating that seat selection has broad (but not universal) appeal.

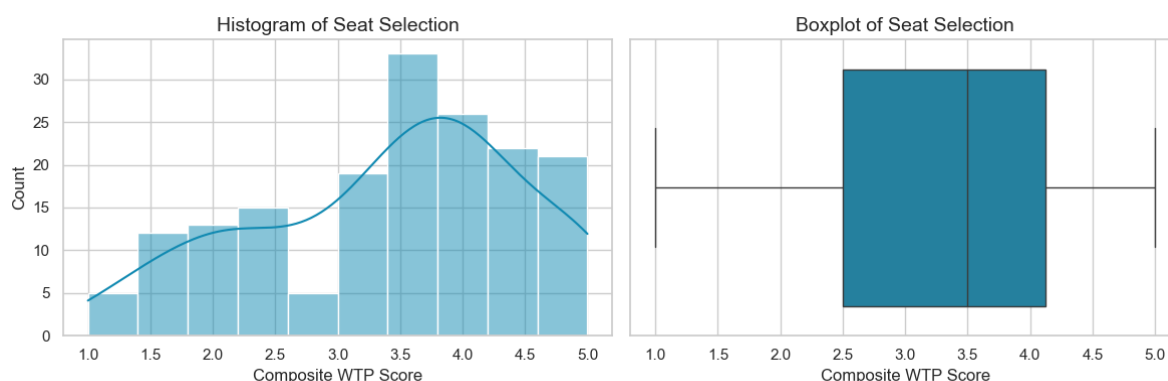


Figure 4.6. Histogram and boxplot of WTP for seat selection.

The distribution for checked baggage (Figure 4.7) is relatively symmetrical and leans slightly lower than for seat selection. The histogram displays a clear central peak and less pronounced tails toward both the lowest and highest scores. This suggests checked baggage is an ancillary service for which willingness to pay is common, but very low or very high WTP is less frequent. The boxplot demonstrates a median around 3.0, and a compact interquartile range, reflecting consistent moderate interest across the sample.

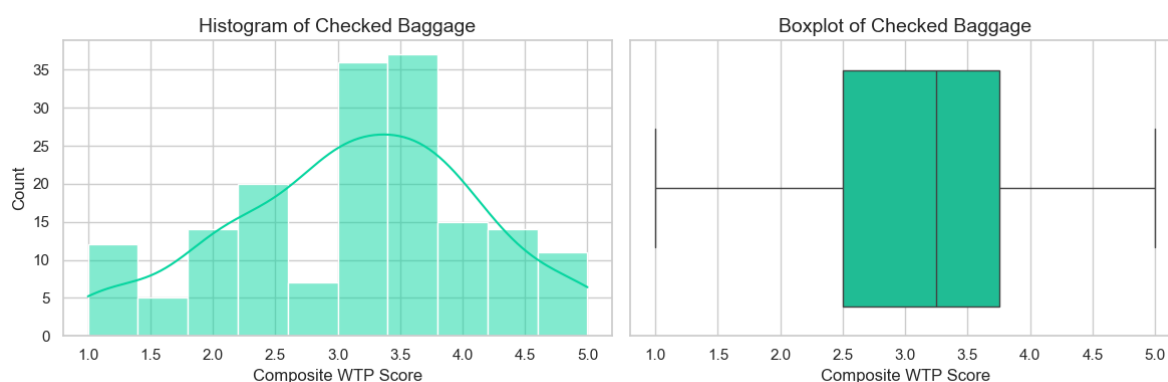


Figure 4.7. Histogram and boxplot of WTP for checked baggage.

Passenger WTP for in-flight meals (Figure 4.8) shows a broader spread and a lower central tendency compared to the previous services. The histogram illustrates a notable accumulation in the lower-to-mid range (around 2.0–3.5), while high willingness to pay is relatively rare. The boxplot confirms this, with a median that appears at 2.5. This suggests in-flight meals may be viewed as a less essential extra, with considerable variability in perceived value among passengers.

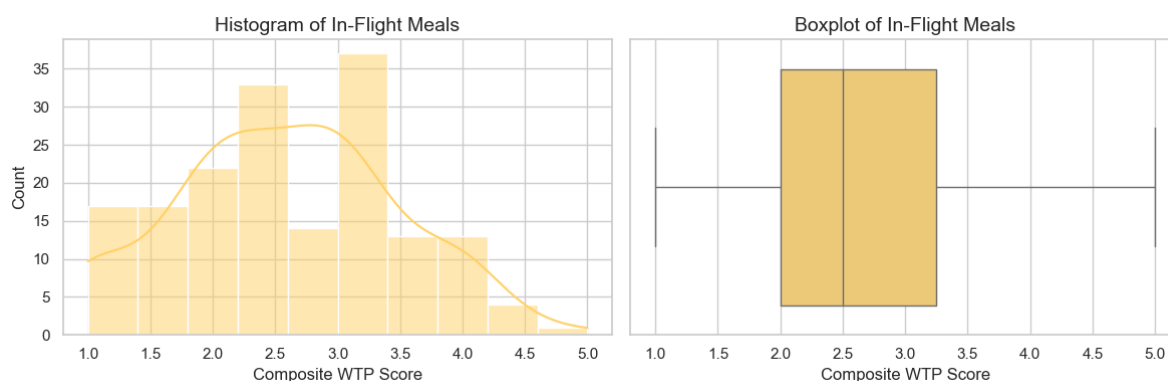


Figure 4.8. Histogram and boxplot of WTP for in-flight meals.

Lastly, the willingness to pay for Wi-Fi (Figure 4.9) shows a clear bimodal pattern. The histogram reveals two distinct peaks: one at the lowest scores (indicating little interest in paying for Wi-Fi), and a second, though slightly smaller, at the highest scores. This split implies polarisation: some respondents see little to no value in in-flight Wi-Fi, while some others view it as highly desirable. The boxplot is correspondingly wide, with the median near the center but substantial spread, highlighting the divergent attitudes within the sample.

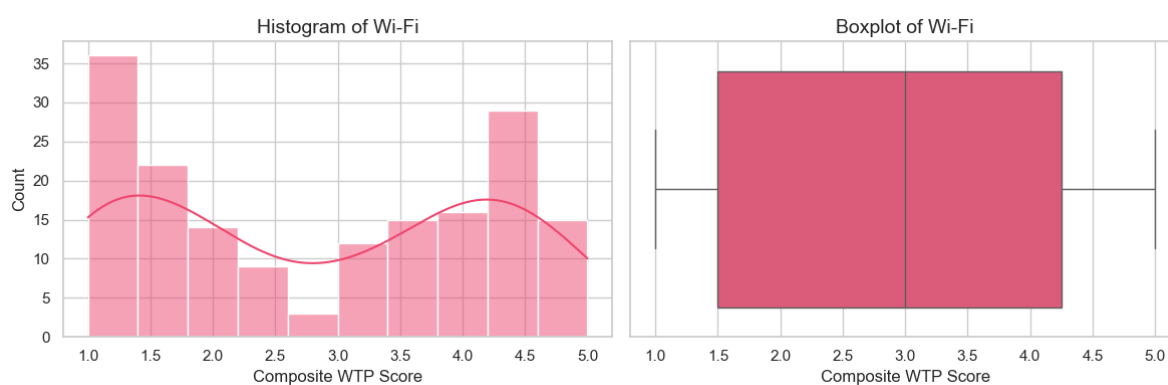


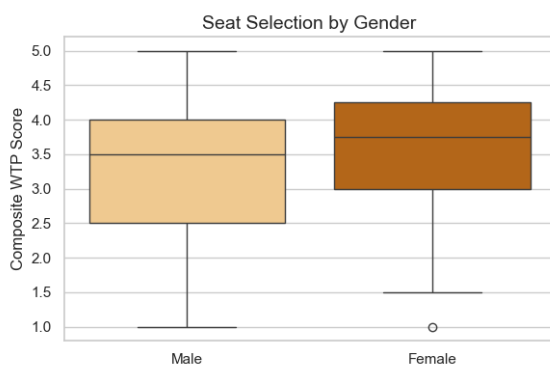
Figure 4.9. Histogram and boxplot of WTP for Wi-Fi.

The visual analysis above reveals significant differences in how ancillary services are valued by air travelers. While seat selection and checked baggage have relatively broad and consistent appeal, in-flight meals and Wi-Fi have a more polarised or generally lower WTP. These findings demonstrate the importance of targeted pricing strategies to capture the full range of passenger preferences.

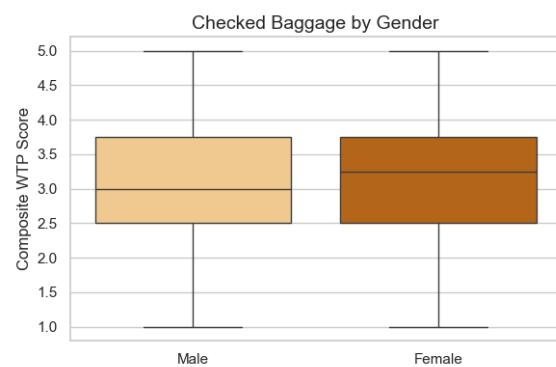
4.1.3 WTP by Passenger Characteristics

Following the overall assessment of WTP for ancillary services, this subsection explores how WTP varies across different passenger characteristics. By analysing how key factors relate to demand for ancillaries, we obtain a better understanding of market segmentation and potential correlations.

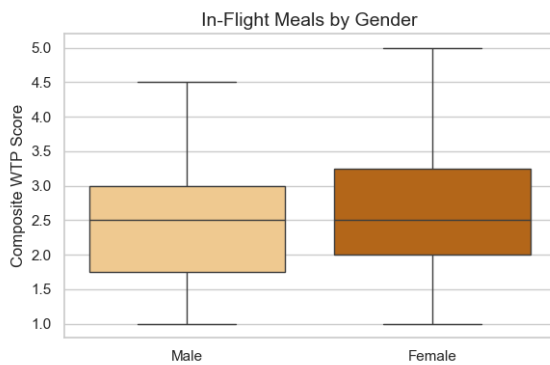
Examining WTP for ancillary services by gender (Figure 4.10) reveals small differences in passenger preferences. For all services, boxplots illustrate subtle but consistent trends. Female respondents tend to have slightly higher median WTP scores than male respondents for all services. The interquartile ranges for both groups overlap substantially; however, the data show that women are slightly more willing to pay for these extras, particularly in the upper quartile.



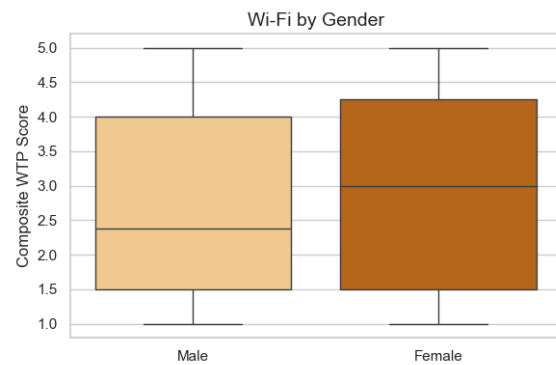
(a) WTP for seat selection by gender.



(b) WTP for checked baggage by gender.



(c) WTP for in-flight meals by gender.



(d) WTP for Wi-Fi by gender.

Figure 4.10. WTP for each ancillary by gender (boxplot).

Age appears as a strong determinant of ancillary service valuation, as seen in the violin plots for each service (Figure 4.11). A clear positive association exists between age and WTP: older age groups tend to report higher composite scores for seat selection, checked baggage, and in-flight meals, while the Wi-Fi service follows a different pattern. These findings show age as a key factor in ancillary service demand segmentation.

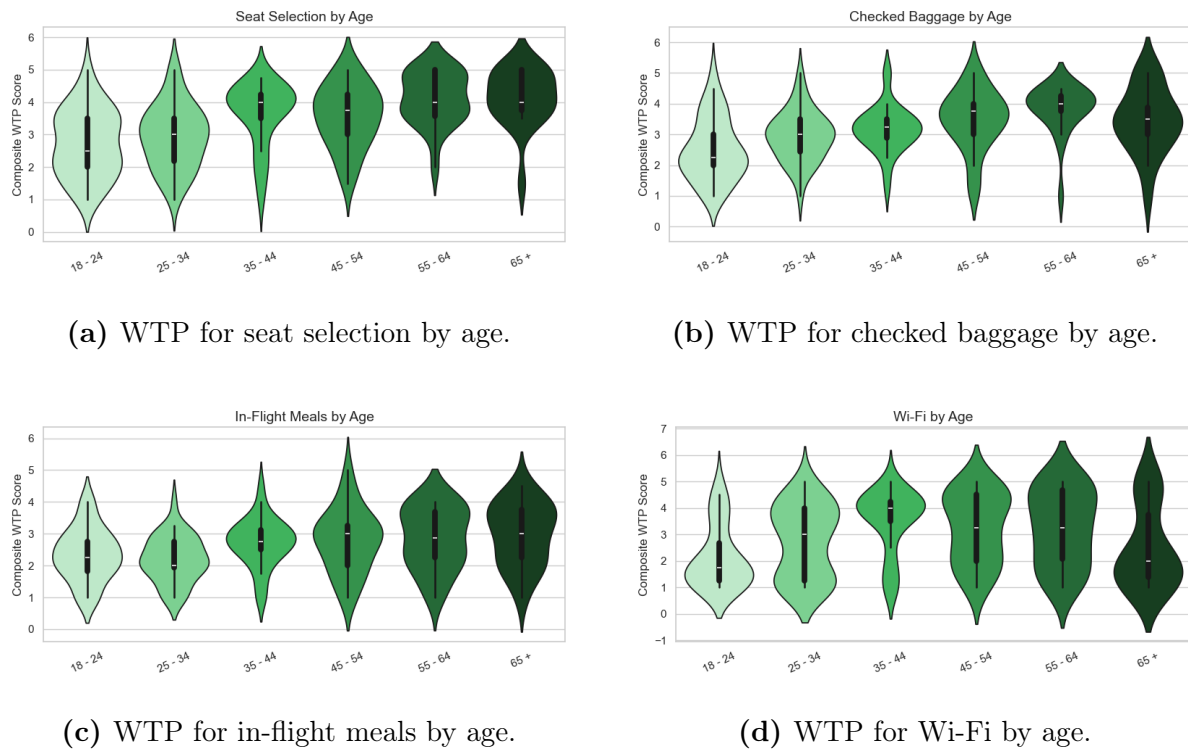


Figure 4.11. WTP for each ancillary by age (violin plot).

Differences in travel purpose (Figure 4.12) also influence variation in WTP. Bar plots show that business travellers have significantly higher mean WTP scores for all ancillary services compared to leisure travellers. This trend is particularly pronounced for seat selection and Wi-Fi, where the business segment's mean is appreciably above that of the leisure segment. These results align with expectations: business passengers often prioritise flexibility, comfort, and connectivity, making them more willing to invest in premium ancillary products.

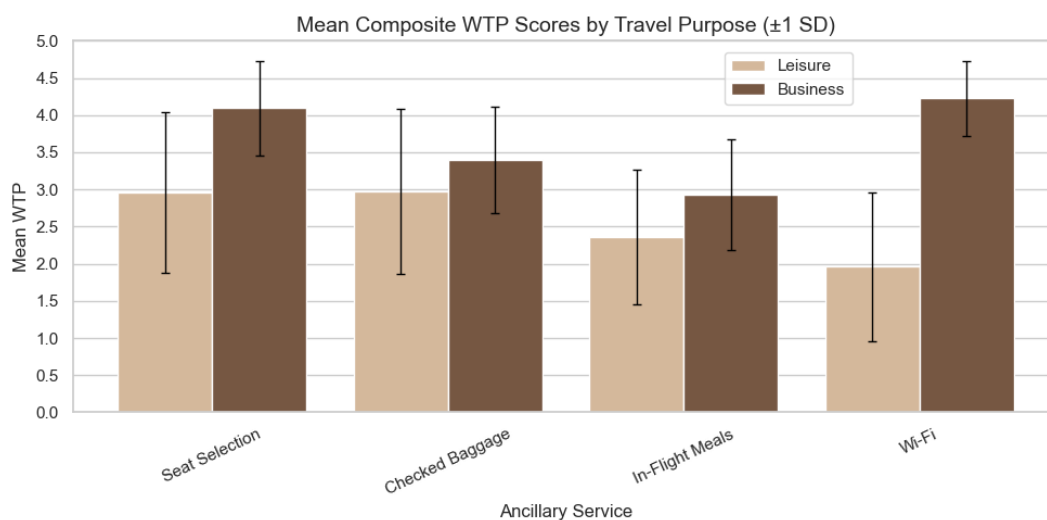


Figure 4.12. Mean WTP for each ancillary service by usual travel purpose (bar plot).

Annual flight frequency (Figure 4.13) offers further information into ancillary purchasing behaviour. The grouped bar plots show that WTP for seat selection remains relatively stable across frequency groups, while checked baggage and in-flight meals show a decreasing trend among more frequent flyers. In contrast, WTP for Wi-Fi increases with annual flight frequency, reflecting increased demand for connectivity among frequent travellers.

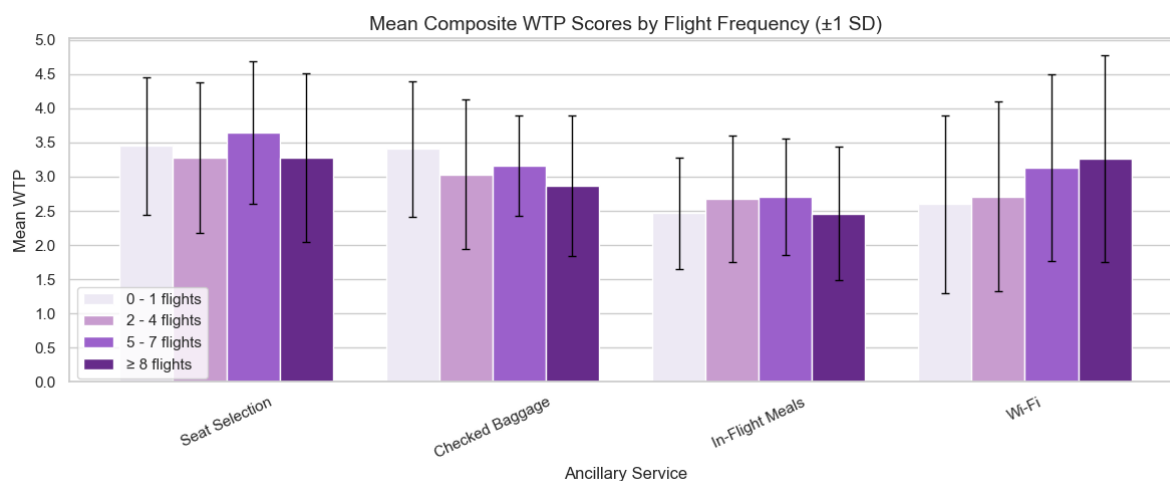


Figure 4.13. Mean WTP for each ancillary service by annual flight frequency (bar plot).

Taken together, these results confirm that passenger characteristics significantly influence ancillary service valuation. Observing these clear profiles justifies deeper, multivariate exploration using machine learning (as presented in the following section), what will enable us to better predict and quantify the relationships between personal attributes and willingness to pay for airline ancillaries.

4.2 Model Performance

This section presents a comparative evaluation of several machine learning algorithms applied to the task of predicting passenger WTP for ancillary services on the BCN-LGW route. The analysis covers four ancillary categories (seat selection, checked baggage, in-flight meals, and Wi-Fi) using four widely adopted classification models: Logistic Regression, Random Forest, SVM, and XGBoost.

To ensure a fair and transparent comparison, all models were trained and evaluated using their default hyperparameters. This approach provides an initial, exploratory baseline that allows for direct comparison of each algorithm's inherent strengths and weaknesses relative to the specific data and target variables. Employing default parameters at this stage helps to:

- Level the playing field across algorithms, highlighting the natural fit of different models to this multiclass classification problem.
- Identify which models are intrinsically well-suited to predicting passenger willingness to pay.
- Detect potential issues related to data preprocessing or feature representation before investing resources to intensive hyperparameter optimisation and model fine-tuning.

All models were therefore evaluated "out of the box" for each ancillary service, with performance assessed using a complete set of metrics: overall accuracy, macro-averaged precision, recall, F1-score, and ROC-AUC. These results establish a clear benchmark for model selection, and provide information into the types of errors and potential improvements to target in future research.

The following subsections present the results organised by ancillary service, summarise the best-performing models, and provide an error analysis via confusion matrices to reveal the specific prediction strengths and weaknesses of each approach.

4.2.1 Performance Metrics by Ancillary Service

This analysis presents the predictive performance of each algorithm, with separate comparisons for every ancillary category.

For seat selection (Table 4.2), Logistic Regression achieved the best overall performance, with the highest F1 score (0.522) and ROC-AUC (0.809). SVM performed second best according to F1 and ROC-AUC, while Random Forest and XGBoost showed comparatively lower values in all metrics. This suggests that, for seat selection, a simpler linear model may be better suited to the structure of the available data, possibly due to cleaner linear boundaries between classes.

SEAT SELECTION					
Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC
Logistic Regression	0.600	0.507	0.569	0.522	0.809
Random Forest	0.486	0.466	0.455	0.459	0.716
SVM	0.543	0.472	0.512	0.483	0.765
XGBoost	0.486	0.415	0.438	0.424	0.664

Table 4.2. Performance metrics for each model predicting WTP for seat selection. The highest value for each metric is highlighted.

For checked baggage (Table 4.3), SVM achieved the highest F1 score (0.711), accuracy (0.714), and precision, indicating both high overall discrimination and balanced class-wise performance. Random Forest performed nearly as well, with slightly lower F1 (0.659) and accuracy. Logistic Regression and XGBoost were less effective on this task. The strong ROC-AUC for Random Forest (0.809) suggests good discrimination even if class-specific prediction was marginally better for SVM.

CHECKED BAGGAGE					
Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC
Logistic Regression	0.629	0.628	0.636	0.621	0.744
Random Forest	0.657	0.668	0.661	0.659	0.809
SVM	0.714	0.728	0.718	0.711	0.782
XGBoost	0.600	0.601	0.607	0.603	0.755

Table 4.3. Performance metrics for each model predicting WTP for checked baggage

For predicting WTP for in-flight meals (Table 4.4), XGBoost outperformed the other models in all metrics except ROC-AUC, with the highest accuracy (0.629), precision (0.612), recall (0.550), and F1 score (0.568); followed by Logistic Regression. SVM's recall and ROC-AUC were strong, but its F1 was lower, likely due to poor precision. This implies that the complex, possibly nonlinear structure of passenger preferences can be better captured by boosting methods for this ancillary, likely by learning subtle interactions in the survey data.

IN-FLIGHT MEALS					
Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC
Logistic Regression	0.600	0.522	0.502	0.504	0.670
Random Forest	0.571	0.472	0.456	0.455	0.687
SVM	0.571	0.356	0.417	0.383	0.760
XGBoost	0.629	0.612	0.550	0.568	0.696

Table 4.4. Performance metrics for each model predicting WTP for in-flight meals

For Wi-Fi WTP (Table 4.5), results are more closely matched. XGBoost had the numerically highest precision (0.655) and F1 score (0.642), while Random Forest demonstrated the highest ROC-AUC (0.899). On the other hand, Logistic Regression and SVM tied for highest accuracy and shared the same F1 score (both 0.829 and 0.595). The relative similarity in these values indicates robust, generalisable prediction for Wi-Fi, possibly because passenger WTP drivers for in-flight Internet are well captured by the available features in most models.

WI-FI					
Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC
Logistic Regression	0.829	0.556	0.643	0.595	0.874
Random Forest	0.800	0.533	0.622	0.574	0.899
SVM	0.829	0.556	0.643	0.595	0.835
XGBoost	0.771	0.655	0.647	0.642	0.825

Table 4.5. Performance metrics for each model predicting WTP for Wi-Fi

As has been seen, performance varied notably by ancillary and model. For seat selection, linear models such as Logistic Regression offered the highest accuracy and F1. For checked baggage, SVM stood out, while XGBoost led for in-flight meals. For Wi-Fi, all algorithms performed well, but XGBoost did it slightly better. These results reinforce the importance of adjusting the algorithm choice to the nature of each ancillary’s data and underlying WTP patterns.

4.2.2 Summary of Best Models

To offer a clear picture of model selection for each ancillary service, Table 4.6 summarises the best-performing algorithm identified for each prediction task, along with its corresponding metrics. The best model for each ancillary was chosen based primarily on the highest macro-averaged F1 score, which balances precision and recall across the three WTP classes and accounts for class imbalance.

As can be seen, model performance varied considerably across ancillary services. For seat selection, Logistic Regression’s F1 score and ROC-AUC suggest that relationships between input features and willingness to pay are largely linear, benefiting from the simplicity and stability of this approach. In contrast, for checked baggage, the SVM significantly outperformed other algorithms, achieving the highest accuracy (0.714) and F1 score (0.711), reflecting its effectiveness in capturing more complex, possibly non-linear boundaries in the data.

Finally, XGBoost appeared as the optimal model for both in-flight meals and Wi-Fi. XGBoost’s ability to model subtle feature interactions likely maximises prediction accuracy in these categories, as reflected in its elevated F1 and ROC-AUC scores. Particularly for Wi-Fi, XGBoost’s ROC-AUC (0.825) surpassed the strong baselines set by the other algorithms, while also demonstrating robust F1 score (0.642).

Ancillary	Best Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC
SS	Log. Reg.	0.600	0.507	0.569	0.522	0.809
CB	SVM	0.714	0.728	0.718	0.711	0.782
IM	XGBoost	0.629	0.612	0.550	0.568	0.696
WF	XGBoost	0.771	0.655	0.647	0.642	0.825

Table 4.6. Summary of the best-performing model for each ancillary service based on macro F1 score, including the main performance metrics.

The F1 score heatmap in Figure 4.14 reinforces these findings visually. Warm colours in the map reveal which algorithm performs best for each service, and help to highlight both strengths and relative weaknesses across the model spectrum. For example, SVM’s standout performance for checked baggage is easily distinguishable. XGBoost, on the other hand, shows particularly strong results for in-flight meals, outperforming the other algorithms for this category. But, in the case of Wi-Fi, while XGBoost achieved the highest F1 score, all four models delivered very similar values, suggesting that the classification task for this service is less challenging and that any of the tested algorithms can provide comparably high predictive accuracy.

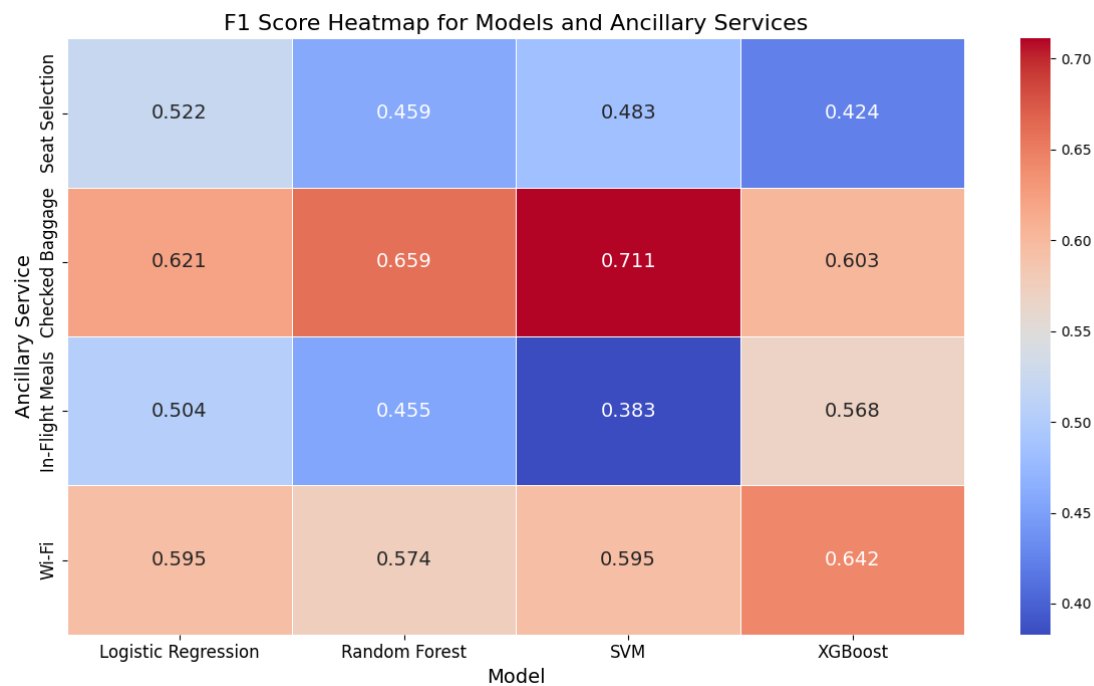


Figure 4.14. Heatmap of macro-averaged F1 scores for each model (columns) and ancillary service (rows).

Overall, these results suggest that no single algorithm is optimal for all ancillaries; instead, each prediction task may benefit from a tailored modelling strategy. However, it is important to remember that these results are based on models trained and evaluated with default hyperparameters. As such, the observed performance rankings are indicative of each algorithm’s baseline fit to the problem, rather than their full potential after tuning. Once hyperparameter optimisation is performed for each model, it is possible that the relative rankings could change, and a model that was not initially the top performer may ultimately achieve superior performance for a given ancillary service. Therefore, these findings should be viewed as a first step in establishing a basis for comparison and guiding future research to optimise and refine predictive models specific to each ancillary context.

4.2.3 Error Analysis: Confusion Matrices

The confusion matrices of the best-performing model for each ancillary service, which were selected earlier, have also been analysed to provide a detailed error assessment. These matrices reveal how effectively the models differentiate between low, medium, and high WTP, highlighting specific misclassification patterns.

The confusion matrix for seat selection using Logistic Regression (Figure 4.15) shows that the majority of "low" and "high" WTP responses are classified correctly, as demonstrated by the higher values along the main diagonal. However, there is a noticeable tendency for the model to confuse "medium" WTP cases with either "low" or "high" classes, which may indicate that responses in the middle range are less distinct and harder to separate on the basis of the available features. Overall, the model demonstrates reasonable discriminatory power, with most errors involving misclassification among adjacent (neighboring) categories.



Figure 4.15. Confusion matrix for the XGBoost model on seat selection.

For checked baggage, the SVM model's confusion matrix (Figure 4.16) reveals that the "low" and "medium" WTP categories are predominantly classified correctly, demonstrating strong discrimination for these groups. However, the model struggles with the "high" WTP class: a significant proportion of high WTP cases are misclassified, particularly as "low." This is an important issue, as it suggests the model is less effective at identifying passengers most willing to pay for checked baggage, which could limit its application in targeted ancillary marketing.

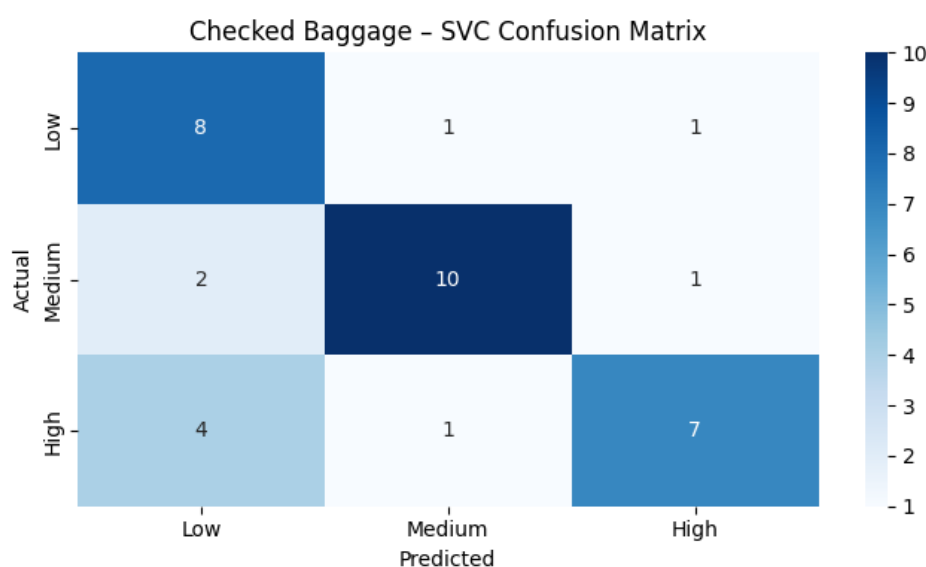


Figure 4.16. Confusion matrix for the XGBoost model on checked baggage.

The XGBoost model for in-flight meals (Figure 4.17) demonstrates strong predictive accuracy for "low" WTP cases, which are well classified. However, about half of the "medium" WTP responses are misclassified (most frequently as "low") while approximately half of the "high" WTP cases are predicted as "medium." This indicates that, while the model easily identifies those not interested in paying extra for in-flight meals, it has greater difficulty distinguishing between genuinely moderate and high willingness among passengers, likely due to less distinct feature separation in this category.

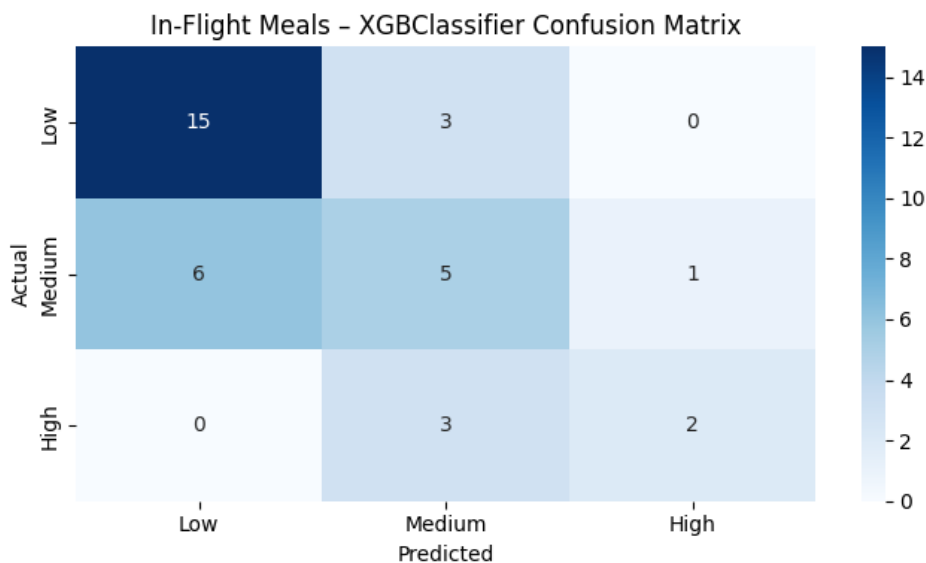


Figure 4.17. Confusion matrix for the SVM model on in-flight meals.

For Wi-Fi, the XGBoost model’s confusion matrix (Figure 4.18) reflects very good overall performance, with the "low" and "high" WTP categories mostly predicted correctly. The primary source of misclassification involves the "medium" WTP group, which tends to be misclassified as either "low" or "high." Despite these errors, the matrix overall is strongly centred on the diagonal, indicating reliable classification for the majority classes, and supporting earlier findings of robust model accuracy for this ancillary.

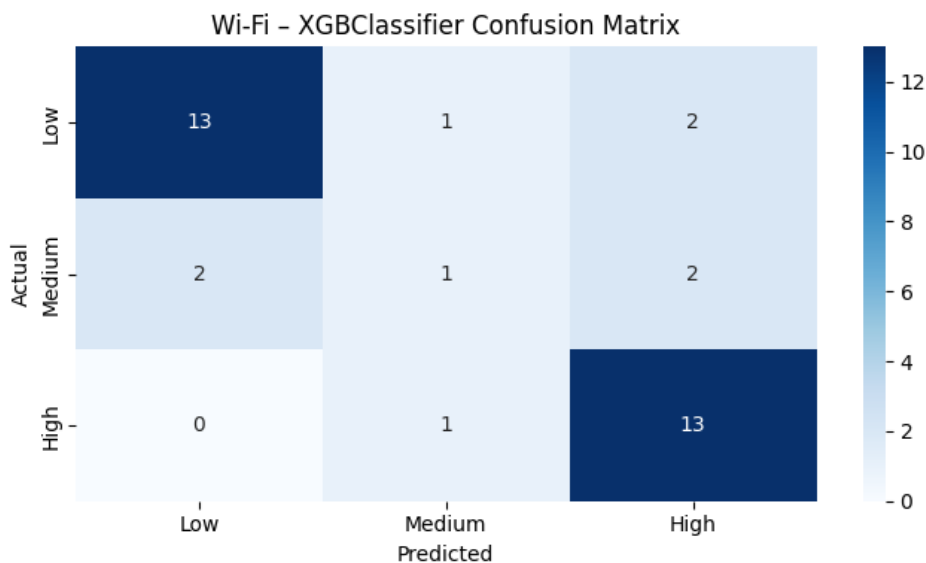


Figure 4.18. Confusion matrix for the Logistic Regression model on Wi-Fi.

In summary, the confusion matrices show that the "medium" WTP group presents the greatest classification challenge across ancillaries, with errors often involving confusion with adjacent classes. These findings point to the importance of focusing on improving differentiation among more ambiguous middle-category passengers and possibly enriching feature selection to better capture their distinct characteristics.

4.3 Feature Importance Analysis

With the predictive models established for each ancillary service, a natural progression is to study which passenger characteristics most influence individual WTP, as understanding feature importance not only improves model interpretability but also provides useful information for adapting ancillary offerings. To address this, two complementary methods were adopted, SHAP values and permutation importance. SHAP analysis offers a granular view of how each input variable shapes model predictions at an individual level, while permutation importance quantifies the overall contribution of each feature to predictive accuracy. This combined approach facilitates a robust identification and ranking of key passenger attributes associated with WTP for each ancillary service.

Analysis of the seat selection ancillary service reveals that passenger age and usual travel purpose are the most influential factors in predicting willingness to pay. The SHAP summary plot for seat selection (Figure 4.19) shows that both Age and Usual Travel Purpose exhibit the greatest spread of SHAP values, indicating these features have the most substantial impact on model output.

Permutation importance scoring (Figure 4.20) reinforces this pattern: Age achieves the highest importance (0.6270), followed by Usual Travel Purpose (0.4531). Both features contribute significantly to the model's predictive accuracy, with shuffling either resulting in a marked decrease in performance. Gender (0.2139) and Annual Flight Frequency (0.2032) provide moderate contributions, while Annual Income (0.0983) plays only a minor role in forecasting a passenger's WTP for seat selection.

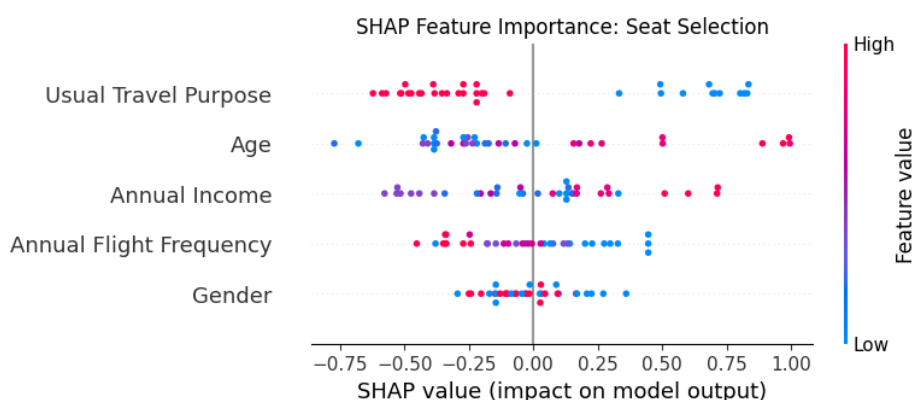


Figure 4.19. SHAP plot of feature impact on predicting seat selection WTP.

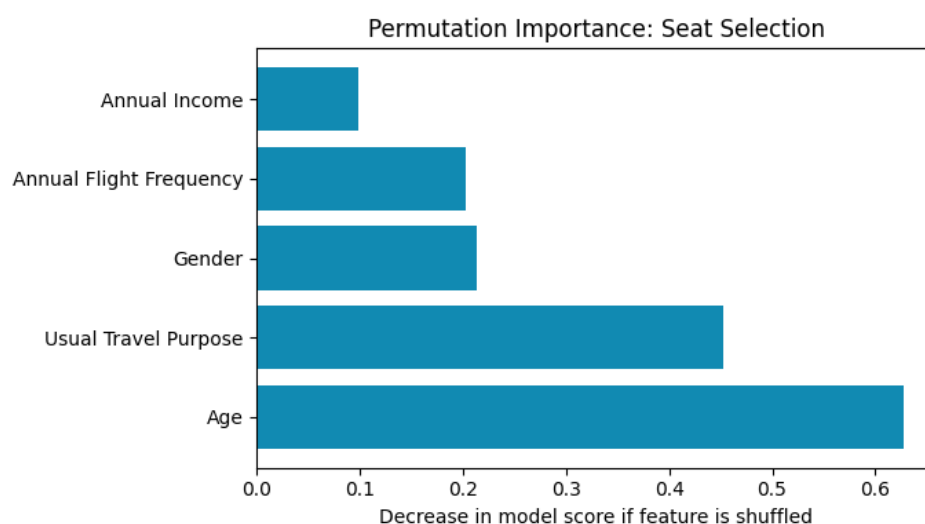


Figure 4.20. Permutation importance of passenger traits for seat selection WTP.

For the checked baggage ancillary service, annual income and age emerge as the most influential predictors of WTP. The SHAP summary plot (Figure 4.21) indicates that Annual Income has a pronounced impact on model predictions, with higher SHAP values demonstrating that variations in income significantly affect the predicted WTP for checked baggage. Age also shows a wide distribution of SHAP values, highlighting its strong individual effect on prediction outcomes.

The permutation importance analysis (Figure 4.22) strongly reinforces these findings. Annual Income achieves the highest importance score (0.7501), indicating that the model's performance declines substantially if this feature is shuffled, while Age follows closely with a score of 0.6337. Both features are thus critical for accurately forecasting whether a passenger is likely to pay for checked baggage. Annual Flight Frequency ranks next (0.3968), suggesting that frequent flyers are more or less inclined to purchase checked baggage based on their typical travel habits. In contrast, Usual Travel Purpose (0.2011) and Gender (0.0644) have comparatively minor effects, making only marginal contributions to the model's overall accuracy for this service.

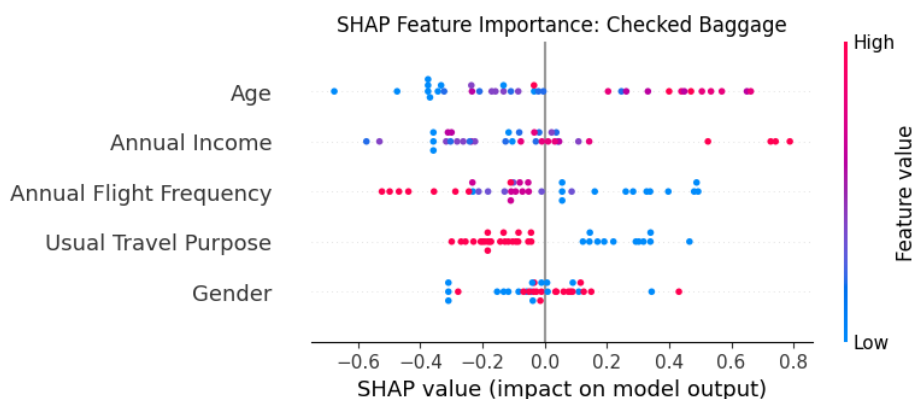


Figure 4.21. SHAP plot of feature impact on predicting checked baggage WTP.

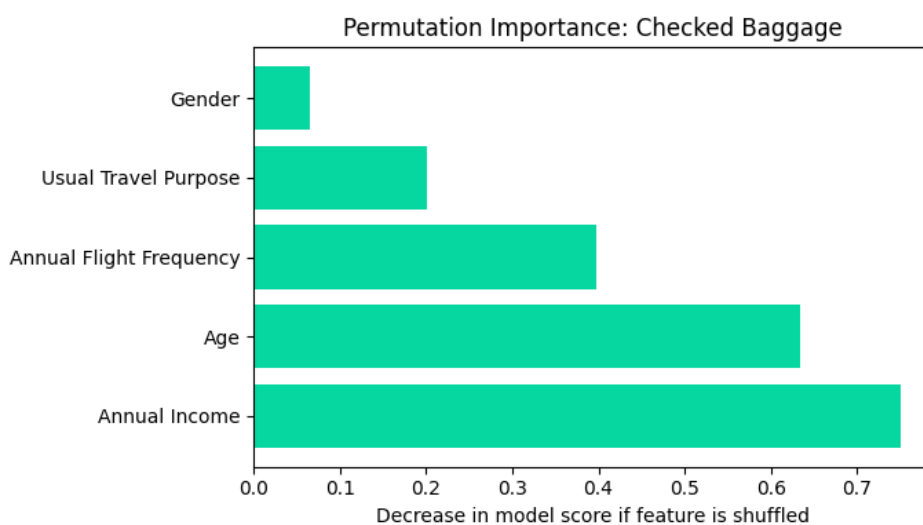


Figure 4.22. Permutation importance of passenger traits for checked baggage WTP.

For the in-flight meals ancillary service, annual income and usual travel purpose are the strongest predictors of passenger WTP. As visualised in the SHAP summary plot (Figure 4.23), Annual Income consistently exhibits the greatest influence on the model’s output, with higher or lower incomes leading to subtle shifts in predicted WTP. Usual Travel Purpose also displays a significant range of SHAP values, highlighting that the reason for travel meaningfully shapes the model’s predictions for this service.

The permutation importance analysis (Figure 4.24) confirms this interpretation: Annual Income holds the highest importance score (0.4501), followed by Usual Travel Purpose (0.3016). These two features play important roles in maintaining the model’s predictive accuracy, as shuffling either results in a notable reduction in model performance. Age, with a much lower score (0.1460), has a more modest effect, while Annual Flight Frequency (-0.0653) and Gender (-0.1396) have negative or essentially negligible importance. These small or negative values indicate that, for in-flight meals, frequently flying and gender do not provide useful discriminative information to the model.

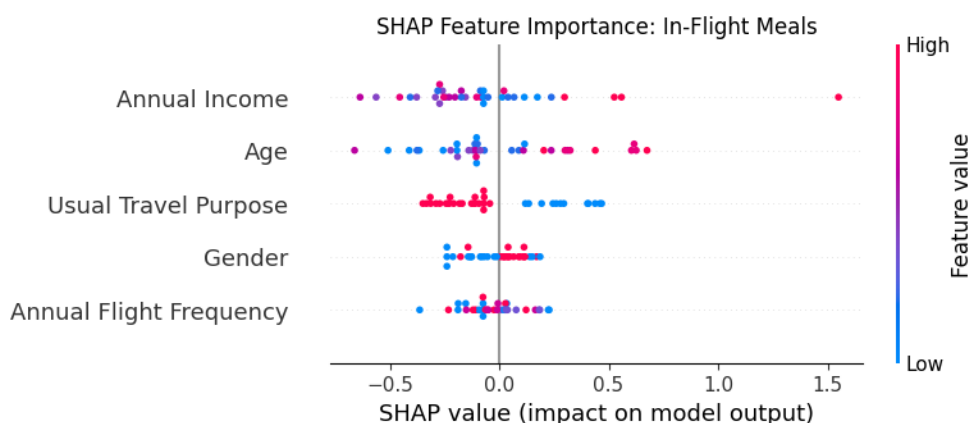


Figure 4.23. SHAP plot of feature impact on predicting in-flight meals WTP.

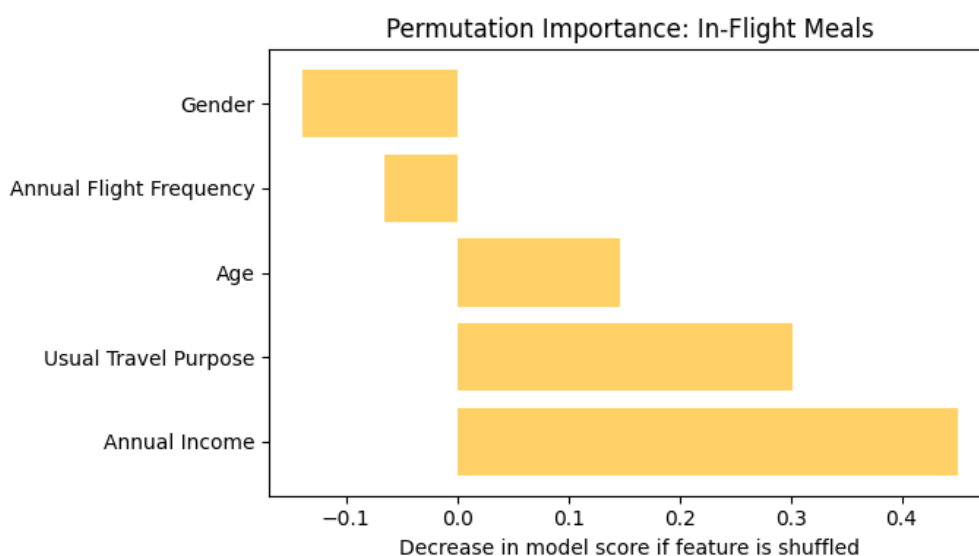


Figure 4.24. Permutation importance of passenger traits for in-flight meals WTP.

For the Wi-Fi ancillary service, the analyses reveal a particularly strong effect of usual travel purpose on WTP. The SHAP summary plot (Figure 4.25) clearly shows that Usual Travel Purpose stands out, with a wide dispersion of SHAP values indicating this feature produces the most substantial and individualised influence on passenger predictions. This suggests that whether a passenger is travelling for business, leisure, or another reason has a major impact on their likelihood of paying for Wi-Fi.

Permutation importance scores (Figure 4.26) reinforce this finding: Usual Travel Purpose records the highest importance by a substantial margin (1.3869), proving its dominant role in model performance, as none of the other features approach this level of influence. Annual Income comes a distant second (0.2910), indicating a moderate effect, while Age (0.0873), Annual Flight Frequency (0.0612), and Gender (-0.0134) contribute little or nothing to the model’s predictive power for this service.

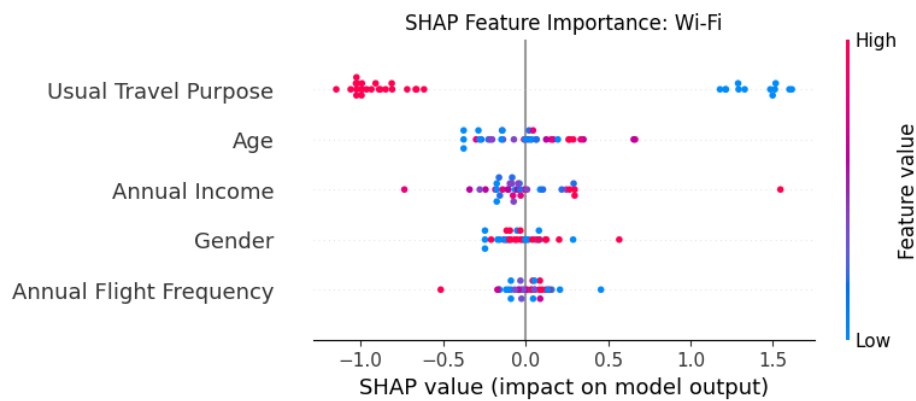


Figure 4.25. SHAP plot of feature impact on predicting Wi-Fi WTP.

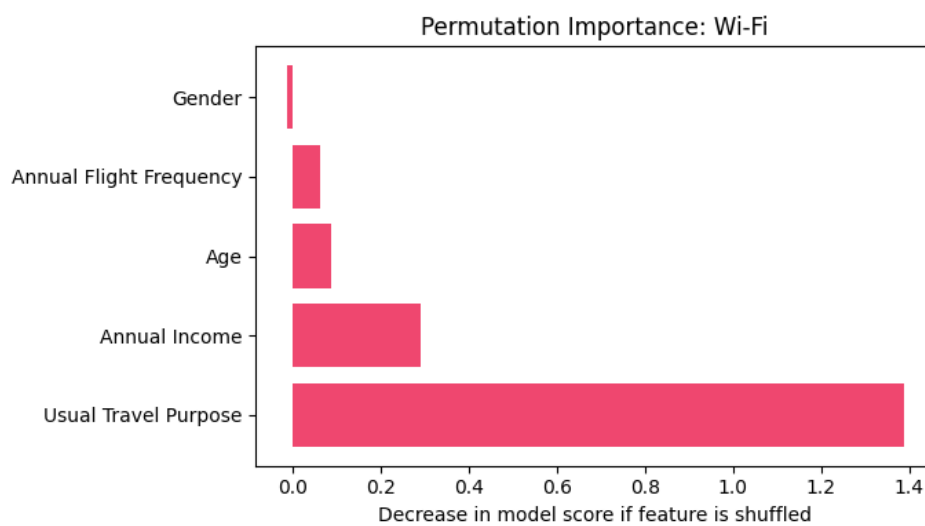


Figure 4.26. Permutation importance of passenger traits for Wi-Fi WTP.

Once the specific results for each ancillary service have been established, it can be interesting to synthesise these findings in a comparative overview (Figure 4.27).

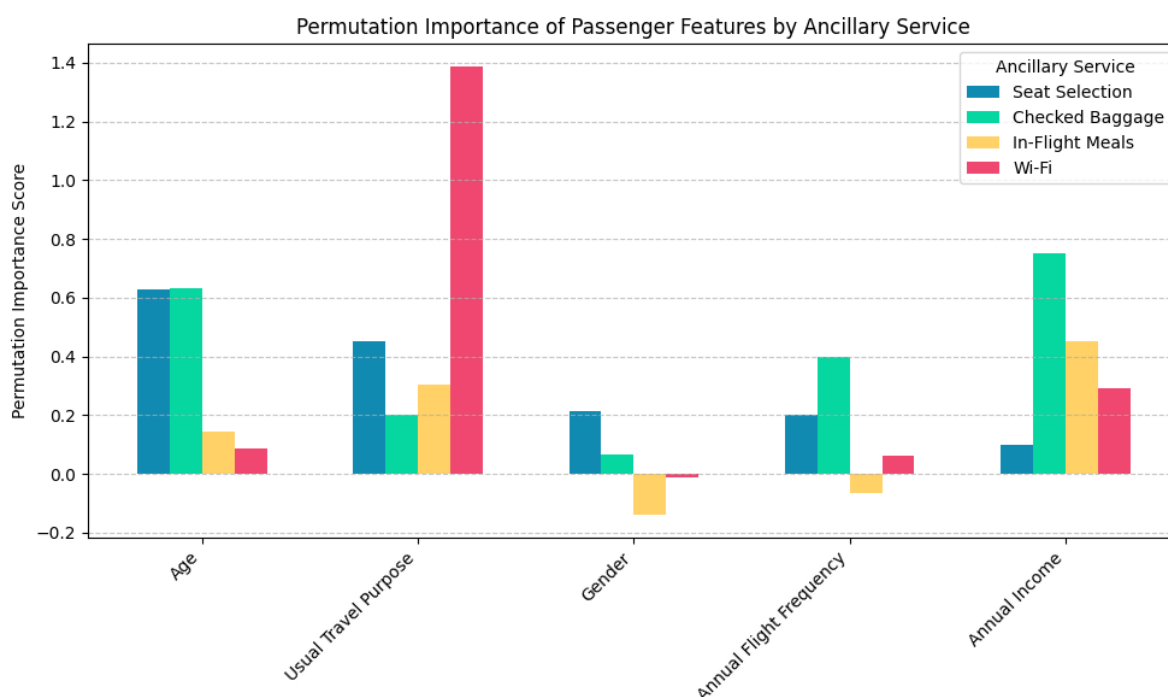


Figure 4.27. Permutation importance of passenger traits for all ancillary services.

This trait-centred perspective reveals that not all passenger characteristics impact ancillary services equally. Most notably, usual travel purpose stands out as the strongest overall predictor, especially for Wi-Fi purchases. This is likely because business travellers, who depend on connectivity for work, are far more willing to pay for in-flight internet, whereas leisure travellers generally show less interest.

Annual income, by contrast, plays its most significant role in predicting demand for checked baggage and in-flight meals. Passengers with higher incomes appear more comfortable incurring extra costs for convenience and comfort, suggesting that discretionary spending power shapes these two purchasing decisions. For services like seat selection or Wi-Fi, however, the effect of income is less pronounced.

Age is another trait that matters most for seat selection and checked baggage. Older travellers may place greater value on certainty, convenience, and comfort, leading to a stronger willingness to pay for a designated seat or additional baggage. In contrast, age has only a minor influence on preferences for in-flight meals and Wi-Fi, which may appeal more uniformly across different age groups.

The frequency of flying moderately affects checked baggage and seat selection, as frequent travellers possibly have different luggage habits or loyalty benefits that change their purchasing behaviour. However, flight frequency has little bearing on Wi-Fi and meal purchases, highlighting that the appeal of these services transcends how often one flies.

Finally, gender emerges as largely irrelevant for all ancillary services, with consistently low importance scores. This suggests that willingness to pay for these extras cannot be effectively predicted by gender.

Altogether, this analysis demonstrates that travel purpose, income, and age are the key traits to consider for targeted ancillary strategies, while gender plays no meaningful role.

Chapter 5

Conclusion

5.1 Interpretation of Findings

This thesis provides a detailed and empirically supported analysis of what influences passenger WTP for ancillary services in the context of the popular short-haul BCN-LGW route. The results show that WTP is strongly differentiated both by the type of ancillary product and by individual passenger characteristics, highlighting the complexity and segmentation of modern air travel markets.

First, a clear hierarchy arises among ancillary products. Seat selection and checked baggage are the most broadly valued extras, with the majority of respondents indicating moderate to high WTP. In particular, demand for seat selection is especially high among older and business travellers, reflecting their desire for greater comfort and predictability. Checked baggage, while still important, has a central but less polarised space, valued by most but rarely generating extreme WTP. In contrast, in-flight meals show much lower and more dispersed WTP, indicating that many passengers perceive them as unnecessary on short-haul flights.

The case of Wi-Fi stands out for its polarisation: a significant group (mostly business travellers and frequent flyers) identifies high value in connectivity, while others would not pay at all. This bimodal distribution suggests significant opportunities for targeted product offerings, but warns against blanket pricing.

Beyond product differences, passenger heterogeneity is also a decisive influence. Age and travel purpose are the strongest drivers: older passengers and those travelling for business are far more likely to exhibit high WTP for seat selection and, in the case of business travellers, for Wi-Fi as well. Income has a significant impact too, especially for checked baggage and in-flight meals; higher-income individuals are far more willing to pay for comfort and convenience, emphasising the importance of discretionary purchasing power. The frequency of flying influences slightly these patterns, amplifying the appeal of certain ancillaries for habitual travellers. Gender, however, has negligible predictive value, indicating that gender segmentation is unlikely to produce effective ancillary strategies in this context.

A core contribution of this thesis lies in the application and analysis of ML models to predict and interpret WTP patterns. By systematically benchmarking several algorithms, the study finds that the optimal predictive model is ancillary-dependent. Linear methods like logistic regression are particularly useful for seat selection, where passenger traits and WTP have clearer, monotonic relationships. For ancillary services where relationships are more complex or nonlinear, such as checked baggage, in-flight meals, and Wi-Fi, advanced algorithms like SVM and XGBoost significantly improve prediction accuracy. Importantly, model interpretability tools (specifically SHAP values and permutation importance) clarified the magnitude and direction of influence of variables like income, age, and travel purpose. Finally, error analysis revealed that 'medium' WTP is the most ambiguous class to forecast, emphasising the challenges of segmenting passengers whose attitudes toward ancillaries are less definite.

Directly addressing the research question, the findings reveal that the factors that most significantly influence passenger WTP for ancillary services on the BCN-LGW route are (in order of impact) travel purpose, annual income, age, and secondarily, annual flight frequency. Business travellers and those with higher incomes consistently report greater willingness to purchase ancillaries, while age is especially predictive for seat selection and checked baggage. In contrast, gender plays a minimal role. Predictive models, when developed with these observed characteristics and employing state-of-the-art ML approaches, can accurately forecast passenger WTP to a meaningful extent. However,

accuracy remains highest for passengers with clearly high or low WTP, while forecasting for the moderate segment proves more challenging, highlighting an opportunity for further modelling refinement and richer data collection.

5.2 Implications for Airlines

The results of this study have direct and actionable implications for airline revenue management and ancillary strategies. The strong influence of travel purpose, age, and income on willingness to pay demonstrates that a generic, one-size-fits-all approach to ancillary sales risks leaving significant revenue untapped. So, airlines can achieve higher profit by personalising offers to specific passenger profiles. For example, business travellers (particularly those on short-haul routes) place high value on flexibility and productivity, making them ideal candidates for premium seat selection or in-flight Wi-Fi packages. Similarly, older passengers exhibit a willingness to invest in comfort and certainty, creating opportunities for targeted seat selection and baggage offers.

Building on this segmentation, the observed variation in WTP supports the adoption of dynamic pricing strategies. Ancillaries such as Wi-Fi and seat selection, which have a high average valuation but a significant fluctuation across passenger groups, are good candidates for adaptive pricing models that adjust in real-time based on demand signals and passenger characteristics. So, by aligning prices more closely with each segment's maximum WTP, airlines can capture greater consumer surplus without alienating price-sensitive travellers.

Finally, the patterns identified around in-flight meals highlight a need for product innovation. The lower and more dispersed WTP scores suggest that current offerings may not adequately match the preferences of short-haul passengers. Airlines could respond by redesigning meal options, introducing lighter or more flexible selections, or experimenting with unbundled individual items, all of which could contribute to increased perceived value and consumption.

5.3 Limitations

Despite providing useful information into passenger WTP for airline ancillary services, this study has several limitations that should be considered when interpreting the findings and the extent of their generalisability.

A primary limitation comes from the sampling methodology. Data collection relied on a snowball sampling technique, initially distributed through my personal and academic networks. While this approach successfully generated a sizeable and diverse sample, it also introduces the risk of self-selection bias. Respondents within these circles may not fully represent the broader travelling population, particularly with respect to higher-income individuals or those who travel frequently for business—groups that may possess different attitudes or purchasing power regarding airline ancillaries. Furthermore, the exclusion of certain underrepresented travel purposes, such as “visiting family or friends” and “education”, was necessary due to minimal response rates. Although methodologically justified, this limits the scope of segmentation and means that the results are most applicable to leisure and business travellers only, which may not capture the full spectrum of motivations and behaviours present on the BCN-LGW route or similar short-haul markets.

Another important limitation is the study’s cross-sectional design. As data was collected at a single point in time, the findings offer only a snapshot of passenger preferences and behaviours. Given the rapidly changing context of the airline industry (shaped by factors such as the COVID-19 pandemic, fluctuations in the economy, or evolving airline product offerings) it is possible that WTP for ancillaries may shift over time. The absence of longitudinal data prevents the detection of such trends or cyclical patterns, highlighting the need for caution in extrapolating these results to other periods or changing market conditions.

Additionally, the research relies on self-reported WTP, as measured by survey responses. While this method is standard in consumer research, it is important to recognise the potential impact of hypothetical bias and social desirability effects. Respondents may overestimate or underestimate their true likelihood of purchasing ancillary services, whether

consciously or unconsciously, resulting in a discrepancy between stated intentions and actual purchasing behaviour. This gap is an inherent challenge in survey-based valuation studies and may affect the precision of the estimated demand for specific ancillaries.

From a methodological perspective, the set of passenger characteristics included in the predictive models was necessarily limited by the scope of the survey instrument. While core variables such as age, income, travel purpose, and frequency were captured, additional features (like trip urgency, group size, loyalty program status, booking channel, or historical purchase data) could meaningfully improve model performance and segmentation accuracy. Their omission represents a missed opportunity for even deeper insight into WTP drivers.

Finally, while the application of ML models unlocked valuable patterns in the data, all algorithms were evaluated using default parameter settings without extensive hyperparameter tuning. Although this approach facilitates straightforward comparison and baseline benchmarking, further optimisation could deliver higher predictive accuracy and, in particular, a better class separation for the ambiguous 'medium' WTP segment.

5.4 Future Research Directions

Building on the current study's findings, several possible future research directions can greatly deepen our understanding of ancillary revenue strategies in the airline industry. A central priority should be the replication of this investigation using broader and more heterogeneous datasets. Extending the research to cover additional air routes, other types of airlines, and diverse geographic regions would help clarify whether the patterns observed on the BCN-LGW route persist across different market segments and cultural contexts.

Beyond the expansion of sample diversity, methodological improvements can offer new insights. One important direction is the use of behavioural experiments or the analysis of revealed-preference data, such as actual ancillary purchases from booking systems. By moving beyond self-reported intentions to real-world behaviours, researchers can reduce

hypothetical and social desirability biases, resulting in more accurate estimations of true WTP. Field experiments(such as randomised ancillary offers or A/B price tests) would allow for causal inference regarding the effectiveness of different ancillary product designs, pricing, and marketing tactics.

Additionally, the dynamic nature of passenger preferences warrants longitudinal analyses. Tracking WTP over time could reveal how consumer attitudes and purchasing behaviours shift in response to external events, such as changes in the economic environment, evolving airline offerings, or even shifts in public health policies post-pandemic. Such temporal data would be invaluable for airlines seeking to adapt pricing strategies and ancillary offerings proactively, rather than reactively.

Expanding the set of features and explanatory variables is another promising area for future work. Including variables such as loyalty program status, group versus solo travel, trip urgency, booking channel (such as mobile versus desktop), or even prior purchase behaviours would allow finer segmentation and more accurate predictive modelling. Richer feature sets can help airlines design even more personalised ancillary packages and targeted promotions, moving toward the vision of “segment of one” marketing.

Advancements in ML methodology also offer attractive potential. Testing additional algorithms, fine-tuning hyperparameters, or adopting more powerful deep learning models could further improves the accuracy of WTP predictions. Moreover, using more complex interpretability frameworks and explainable AI tools beyond SHAP and permutation importance would offer more practical information into the drivers of ancillary purchases.

Finally, future studies could explore how the bundling of ancillary services (offering value-added packages rather than à la carte products) shapes both WTP and passenger satisfaction. Understanding the relationship between perceived value, product bundling, and purchasing decisions can help airlines craft ancillary strategies that not only boost revenue, but also improve the overall customer experience.

Bibliography

- Aerospace Global News. (2024). *Vueling reveals gatwick – barcelona as its most popular uk route*. Aerospace Global News. <https://aerospaceglobalnews.com/news/vueling-reveals-gatwick-barcelona-as-its-most-popular-uk-route/>
- Aggregate Intelligence. (2025, May). *How can airlines increase ancillary revenue through ai-powered upselling and cross-selling?* Aggregate Intelligence . <https://aggregateintelligence.com/how-can-airlines-increase-ancillary-revenue-through-ai-powered-upselling-and-cross-selling/>
- AlHabbal, M. R. (2022, December). *Predicting & optimizing airlines customer satisfaction using classification*. Thesis. Rochester Institute of Technology. <https://repository.rit.edu/theses/11383/>
- Altman, D. G., & Royston, P. (2006, May). *The cost of dichotomising continuous variables*. BMJ. <https://www.bmj.com/content/332/7549/1080.1.extract>
- Azad, M. S., Khan, S. S., Hossain, R., Rahman, R., & Momen, S. (2023, December). *Predictive modeling of consumer purchase behavior on social media: Integrating theory of planned behavior and machine learning for actionable insights*. Plos one, 18(12). <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0296336>
- Biau, G., & Scornet, E. (2016). *A random forest guided tour*. Springer - Test. <https://link.springer.com/article/10.1007/s11749-016-0481-7>
- Buchmann, A. (2025, June). *Willingness to pay (wtp) for leisure air travel: From wtp analysis to the conception of the 3-plane interaction model (3-pim) as decision-making tool for airline management*. Elsevier - Journal of the Air Transport Re-

- search Society. <https://www.sciencedirect.com/science/article/pii/S2941198X25000077>
- Chiambareto, P. (2021, March). *Air passengers' willingness to pay for ancillary services on long-haul flights*. Elsevier - Transportation Research Part E: Logistics and Transportation Review. <https://www.sciencedirect.com/science/article/abs/pii/S1366554521000120>
- Cooper, A. (2021, November). *Explaining machine learning models: A non-technical guide to interpreting shap analyses*. Aidan Cooper. <https://www.aidancooper.co.uk/a-non-technical-guide-to-interpreting-shap-analyses/>
- Cordes, H., Baumeister, S., & Käyrä, M. (2024, March). *Factors influencing the willingness to pay for aviation voluntary carbon offsets: A literature review*. European Journal of Tourism Research. <https://www.ejtr.vumk.eu/index.php/about/article/view/2741>
- Dureh, N., Ueranantasan, A., & Eso, M. (2018). *A comparison of multiple linear regression and random forest for community concern of youth and young adults survey*. Methods. https://doi.nrct.go.th/admin/doc/doc_624672.pdf
- El-Hage, S. (2020). *Predicting purchase of airline seating using machine learning*. KTH, School of Electrical Engineering and Computer Science (EECS). <https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1465383&dswid=-2244>
- Field, A. (2024). *Discovering statistics using ibm spss statistics*. Sage publications limited.
- Fletcher, T. (2008). *Support vector machines explained*. UCL. https://www.csd.uwo.ca/~xling/cs860/papers/SVM_Explained.pdf
- Galli, S. (2024, January). *Understanding permutation feature importance for model interpretation*. Train In Data. <https://www.blog.trainindata.com/permutation-feature-importance/>

- GeeksForGeeks. (2025). *Xgboost*. GeeksForGeeks. <https://www.geeksforgeeks.org/machine-learning/xgboost/>
- Hair, J. F. (2014). *Multivariate data analysis*. (7th ed.). Pearson.
- Hamdani, N., Permana, I., Nugraha, S., & Mubarok, T. M. S. (2021, October). *Are passengers willingness to pay extra money for domestic flight low-cost carrier airline in covid-19 pandemic?* Studies of Applied Economics. <https://ojs.ual.es/ojs/index.php/eea/article/view/5354>
- Hinnen, G., Hille, S. L., & Wittmer, A. (2015, November). *Willingness to pay for green products in air travel: Ready for take-off?* Business Strategy and the Environment. <https://onlinelibrary.wiley.com/doi/abs/10.1002/bse.1909>
- Hoang, J. (2023, June). *How to evaluate computer vision models with confusion matrix*. Datature. <https://datature.io/blog/how-to-evaluate-computer-vision-models-with-confusion-matrix>
- Hui, T. H., & O'Connell, J. F. (2024, July). *Examining air travellers' willingness to pay for non-voluntary environment-related fees: The case of saf surcharge and carbon taxes*. Highlights of Sustainability. <https://www.hos.pub/articles/hsustain3010005>
- IATA. (2025, June). *Airline profitability to strengthen slightly in 2025 despite headwinds*. 2025 Press Releases. <https://www.iata.org/en/pressroom/2025-releases/2025-06-02-01/>
- Instituto Nacional de Estadística. (2023). *Distribución salarial. percentiles de salarios. trabajadores en función de su ganancia anual respecto al salario mínimo interprofesional (smi). deciles de salarios. proporción de trabajadores con ganancia baja (low pay rate)*. Wages, income, social cohesion. https://www.ine.es/ss/Satellite?L=en_GB&c=INESeccion_C&cid=1259931351611&p=1254735110672&pagename=ProductosYServicios%2FPYSLayout&param1=PYSDetalle&param3=1259924822888

- Jander, J. (2024). *Breaking down willingness-to-pay in airline rm.* PROS Insights. <https://pros.com/airline-offer-optimization-guide-willingness-to-pay/>
- Jasra, S. K., Valentino, G., Muscat, A., & Camilleri, R. (2025, February). *A comparative study of unsupervised machine learning methods for anomaly detection in flight data: Case studies from real-world flight operations.* MDPI - Aerospace. <https://www.mdpi.com/2226-4310/12/2/151>
- JetMate. (2025, February). *The future of aviation pricing: Ai-powered willingness-to-pay (wtp) models.* JetMate. <https://jetmate.aero/articles/the-future-of-aviation-pricing-ai-powered-willingness-to-pay-wtp-models/>
- Kassir, A., & Ashaal, A. (2021, December). *Ancillary revenue of full service airline an analysis of the passenger's willingness to pay [wtp](the case of middle east airlines).* Review of Politics and Public Policy in Emerging Economies. <https://publishing.globalcsrc.org/ojs/index.php/rope/article/view/2234>
- Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006, January). *Data preprocessing for supervised leaning.* International journal of computer science. https://www.researchgate.net/publication/228084519_Data_Preprocessing_for_Supervised_Learning
- Lee, F. (2020). *What is logistic regression?* IBM. <https://www.ibm.com/think/topics/logistic-regression>
- Lestari, P., Utomo, D., & Nurhidayat, A. (2019, October). *Analysis of ability to pay and willingness to pay of aircraft passengers toward jakarta surabaya high-speed train.* 11th Asia Pacific Transportation and the Environment Conference (APTE 2018). <https://www.atlantis-press.com/proceedings/apte-18/125918731>
- Lin, J. (2025, May). *Application of machine learning in predicting consumer behavior and precision marketing.* PLoS One. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0321854>

- Lindhjem, H., & Navrud, S. (2012, November). *Using internet in stated preference surveys: A review and comparison of survey modes*. International Review of Environmental and Resource Economics (Forthcoming). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1914273
- Lindhjem, H., & Navrud, S. (2019). *The contingent valuation method*. Economic valuation of natural resources. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780429043260-9/contingent-valuation-method-richard-bishop-thomas-heberlein>
- Liu, H., Abdullah, N. H. B., & Lee, S. Y. (2024, February). *A review of ancillary services in the airline industry*. Taylor & Francis - Cogent Business & Management. <https://www.tandfonline.com/doi/full/10.1080/23311975.2024.2322018>
- Lu, Y. (2019). *Impacts of ancillary services on airline revenue management*. Massachusetts Institute of Technology - PhD Thesis. <https://dspace.mit.edu/handle/1721.1/122413>
- Maheshwari, A., Davendralingam, N., & DeLaurentis, D. A. (2018, June). *A comparative study of machine learning techniques for aviation applications*. 2018 Aviation Technology, Integration, and Operations Conference. <https://arc.aiaa.org/doi/abs/10.2514/6.2018-3980>
- Molin, E., Blangé, J., Cats, O., & Chorus, C. (2017, July). *Willingness to pay for safety improvements in passenger air travel*. Elsevier -Journal of Air Transport Management. <https://www.sciencedirect.com/science/article/abs/pii/S0969699716304756>
- Molnar, C. (2019a, February). *Global model-agnostic methods - permutation feature importance*. Interpretable Machine Learning. <https://christophm.github.io/interpretable-ml-book/shap.html>

- Molnar, C. (2019b, February). *Local model-agnostic methods - shap*. Interpretable Machine Learning. <https://christophm.github.io/interpretable-ml-book/shap.html>
- Naidu, G., Zuva, T., & Sibanda, E. M. (2023). *A review of evaluation metrics in machine learning algorithms*. Springer - Computer science on-line conference. https://link.springer.com/chapter/10.1007/978-3-031-35314-7_2
- Nanyonga, A., Wasswa, H., Joiner, K., Turhan, U., & Wild, G. (2024, November). *Explainable supervised learning models for aviation predictions in australia*. MDPI - Aerospace. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4969197
- Norman, G. (2010, February). *Likert scales, levels of measurement and the “laws” of statistics*. Advances in Health Sciences Education. <https://link.springer.com/article/10.1007/S10459-010-9222-Y>
- OAG. (2023, June). *Shaping airline retail: The unstoppable rise of ancillaries*. OAG. <https://www.oag.com/blog/shaping-airline-retail-unstoppable-rise-ancillaries>
- OAG. (2024, July). *Navigating generational shifts in the airline industry*. OAG Blog - Future of Travel. <https://www.oag.com/blog/generational-shifts-airline-industry>
- OAG. (2025). *Oag analyser*. <https://analytics.oag.com/analyser-client/modules>
- Ozmec-Ban, M., Škurla Babić, R., Vidović, A., & Bračić, M. (2022, July). *A review of ancillary services implementation in the revenue management systems*. Promet-Traffic&Transportation. <https://hrcak.srce.hr/clanak/412852>
- Pisner, D. A., & Schnyer, D. M. (2020). *Support vector machine*. Machine learning. <https://www.sciencedirect.com/science/article/abs/pii/B9780128157398000067>

- Puranik, T. G., Rodriguez, N., & Mavris, D. N. (2020, November). *Towards online prediction of safety-critical landing metrics in aviation using supervised machine learning*. Elsevier - Transportation Research Part C: Emerging Technologies. <https://www.sciencedirect.com/science/article/abs/pii/S0968090X20307245>
- Rocheva, A., Varshaver, E., & Ivanova, N. (2022). *Surveys: The challenge of biases*. Springer Nature - Migration Research in a Digitized World: Using Innovative Technology to Tackle Methodological Challenges. https://www.ssoar.info/ssoar/bitstream/handle/document/81868/ssoar-2022-potzschke_et_al-Migration_Research_in_a_Digitized.pdf
- Sanders, M. N. (2023). *Research methods for business students* (9th). Pearson.
- Shahid, R., Mozumder, M. A. S., Sweet, M. M. R., Hasan, M., Alam, M., Rahman, M. A., Prabha, M., Arif, M., Ahmed, M. P., & Islam, M. R. (2024, June). *Predicting customer loyalty in the airline industry: A machine learning approach integrating sentiment analysis and user experience*. International Journal on Computational Engineering. <https://comien.org/index.php/comien/article/view/12>
- Shaw, M., Tiernan, S., O'Connell, J. F., Warnock-Smith, D., & Efthymiou, M. (2021, January). *Third party ancillary revenues in the airline sector: An exploratory study*. Elsevier - Journal of Air Transport Management. <https://www.sciencedirect.com/science/article/abs/pii/S0969699720305196>
- Sheridan Stavac, C. H., & Giannetti, G. B. (2024, June). *Driving growth with airline ancillaries*. MasterCard Services - Consumer engagement & loyalty consulting. <https://www.mastercardservices.com/en/industries/airlines/insights/driving-growth-airline-ancillaries>
- Singh, R. (2025, January). *Why the airline industry is one of the least profitable sectors – and who really profits in the value chain*. Local and Global Eco. <https://localandglobaleco.com/2025/01/02/why-the-airline-industry-is-one->

of-the-least-profitable-sectors-and-who-really-profits-in-the-value-chain/

Škurla Babić, R., Ozmec-Ban, M., & Bajić, J. (2019, May). *A review of recent trends in airline ancillary revenues*. Časopis za ekonomiju i tržišne komunikacije. <https://www.emc-review.com/sites/default/files/2019-1/Ruzica%20Skurla.pdf>

Tathagat, C., & Montanari, S. (2024, November). *Strategies for airlines to improve ancillary sales*. Cognizant. <https://www.cognizant.com/nl/en/insights/blog/articles/strategies-for-airlines-to-improve-ancillary-sales>

Travel Technology & Solutions. (2025). *Importance of ancillary services in the travel industry*. Travel Technology & Solutions. <https://www.tts.com/blog/the-evolution-of-ancillary-services-in-the-travel-industry/>

UK Civil Aviation Authority. (2024). *Flying in 2023*. Aviation Consumer Survey. <https://www.caa.co.uk/publication/download/20986>

UK Government. (2018). *Dynamic surveying for aviation: Business passengers*. Department for Transport. <https://assets.publishing.service.gov.uk/media/5f608dcdd3bf7f7234487bda/dynamic-surveying-business-passengers-document.pdf>

UK Government. (2022, September). *Household income*. Work, pay and benefits. <https://www.ethnicity-facts-figures.service.gov.uk/work-pay-and-benefits/pay-and-income/household-income/latest/>

Urhausen, J. (2008). *Tourism in europe: Does age matter?* EUROSTAT, Office for Official Publications of the European Communities. <https://edz.bib.uni-mannheim.de/edz/pdf/statinf/08/KS-SF-08-069-EN.pdf>

Viasat. (2023, September). *Free in-flight wi-fi now most influential factor, other than ticket price, for airline passengers, reveals new viasat survey*. Viasat. <https://www.viasat.com/news/latest-news/aviation/2023/free-in-flight-wi-fi->

[now-most-influential-factor-other-than-ticket-price-for-airline-passengers-reveals-new-viasat-survey/](#)

Warnock-Smith, D., O'Connell, J. F., & Maleki, M. (2017, September). *An analysis of ongoing trends in airline ancillary revenues*. Elsevier - Journal of Air Transport Management. <https://www.sciencedirect.com/science/article/pii/S2941198X25000077>

Wilson, J. (2025, May). *Analytical modeling of ancillary items*. Analytics. <https://www.mdpi.com/2813-2203/4/2/14>

Workie, H. A., Tesfa, T., & Kassa, L. (2024, November). *In-flight sales prediction using machine learning*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4969197

Xu, B., Ahmad, S., Charles, V., & Xuan, J. (2022, November). *Sustainable commercial aviation: What determines air travellers' willingness to pay more for sustainable aviation fuel?* Elsevier - Journal of cleaner production. <https://www.sciencedirect.com/science/article/pii/S0959652622035624>

Zaidi, A., & Al Luhayb, A. S. M. (2023, April). *Two statistical approaches to justify the use of the logistic function in binary logistic regression*. Wiley Online Library - Mathematical Problems in Engineering. <https://onlinelibrary.wiley.com/doi/10.1155/2023/5525675>

Zhang, P., Jia, Y., & Shang, Y. (2022). *Research and application of xgboost in imbalanced data*. International Journal of Distributed Sensor Networks. <https://journals.sagepub.com/doi/full/10.1177/15501329221106935>

Zhou, Y., Zhang, T., Mo, Y., & Huang, G. (2020, December). *Willingness to pay for economy class seat selection: From a chinese air consumer perspective*. Elsevier - Research in Transportation Business & Management. <https://www.sciencedirect.com/science/article/abs/pii/S2210539519303451>

Zhou, Y., Loi, A. M.-W., Tan, G. W.-H., Lo, P.-S., & Lim, W. (2022, June). *The survey dataset of the influence of theory of planned behaviour on purchase behaviour on social media*. Science Direct - Data in Brief. <https://www.sciencedirect.com/science/article/pii/S2352340922004413>

Appendix A

CURES Approval Letter

As part of Cranfield University's commitment to research integrity and ethical compliance, all research projects (particularly those involving primary data collection) must receive formal approval through the Cranfield University Research Ethics System (CURES). This appendix includes the official CURES approval letter confirming that the study's design, including its survey-based data collection methodology, adheres to the university's Research Ethics Policy and the principles of the Concordat to Support Research Integrity. The approval was granted prior to survey distribution, ensuring compliance with regulatory requirements and ethical best practices, including participant confidentiality.



26 June 2025

Dear Miss Sorribas Bono ,

Reference: CURES/25827/2025

Project ID: 29377

Title: Predicting Passenger Willingness to Pay for Ancillary Services on the BCN-LGW Route: A Machine Learning Approach Using Likert-Scale Survey Data

Thank you for your application to the Cranfield University Research Ethics System (CURES).

We are pleased to inform you your CURES application, reference CURES/25827/2025 has been reviewed. You may now proceed with the research activities you have sought approval for.

If you have any queries, please contact CURES Support.

We wish you every success with your project.

Regards,

CURES Team

Figure A.1. CURES approval letter.

Appendix B

Survey Document

This appendix presents the survey instrument used to collect primary data for the study. The questionnaire, titled "*Passenger Survey: Willingness to Pay for Ancillary Services on the BCN-LGW Route*", was designed to assess passenger preferences for add-on services (seat selection, checked baggage, in-flight meals, and Wi-Fi) offered by low-cost carriers on short-haul flights. Administered electronically, the survey emphasised anonymity and voluntary participation, with a scenario-based introduction to contextualise responses.

Passenger Survey: Willingness to Pay for Ancillary Services on the BCN-LGW Route

Scenario:

Imagine you are planning to travel on a short-haul flight between Barcelona (BCN) and London Gatwick (LGW) with a low-cost airline, such as Vueling or Ryanair. As is common with these airlines, the base ticket price covers only your seat on the plane, and additional services—such as seat selection, checked baggage, in-flight meals, or Wi-Fi—are offered separately for an extra fee.

When responding to the following questions, please base your answers on your own experiences and preferences as a passenger on this type of flight. Consider your past journeys on similar low-cost carriers, and answer as honestly as possible about what matters most to you when booking and flying this route.

There are no wrong responses to any of these statements, and your participation in this survey is entirely voluntary and anonymous. You are free to stop the survey at any time before submission by closing the browser or exiting the page. However, once you click ‘Submit,’ your responses cannot be withdrawn as all data is anonymized and untraceable. By proceeding, you confirm your consent to participate under these terms.

Section A: Demographic & Travel Profile

Please provide the following details about yourself by checking the appropriate boxes:

1. Gender:

- ☐ Male ☐ Female ☐ Prefer not to say

2. Age:

- ☐ 18–24 ☐ 35–44 ☐ 55–64
☐ 25–34 ☐ 45–54 ☐ 65+

3. Annual income (€):

- ☐ <20,000 ☐ 35,000–49,999 ☐ 75,000–99,999
☐ 20,000–34,999 ☐ 50,000–74,999 ☐ ≥100,000

4. Usual travel purpose:

- ☐ Leisure ☐ Visiting family/friends
☐ Business ☐ Education

5. Flights per year:

- ☐ 0–1 flights ☐ 5–7 flights
☐ 2–4 flights ☐ ≥8 flights

Figure B.1. Page 1 of the passenger survey.

Section B: Willingness to Pay for Ancillary Services

For each statement below, please indicate your level of agreement:

No	Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Seat Selection (SS)						
SS1	Having the ability to choose my seat in advance makes me feel more in control of my journey	1	2	3	4	5
SS2	Sitting in a preferred seat (window, aisle, or extra legroom) improves my overall flight experience	1	2	3	4	5
SS3	I consider having my preferred seat important, even if it means spending less on other aspects of my trip	1	2	3	4	5
Checked Baggage (CB)						
CB1	Having the option to check a bag is important for my travel needs	1	2	3	4	5
CB2	I feel more relaxed when I know my checked baggage is included in my booking	1	2	3	4	5
CB3	I prefer to pay for checked baggage rather than limit myself to only cabin baggage to save money	1	2	3	4	5
In-Flight Meals (IM)*						
IM1	The availability of meal options on board contributes to a more comfortable flight	1	2	3	4	5
IM2	Having a meal on board is important enough to me that I am likely to pay for it, even on short flights	1	2	3	4	5
IM3	Quality meal offerings on board influence my perception of the airline	1	2	3	4	5
Wi-Fi (WF)						
WF1	Access to Wi-Fi during the flight is important for me	1	2	3	4	5
WF2	I prefer to pay for in-flight Wi-Fi rather than fly without internet access to save money	1	2	3	4	5
WF3	The availability of in-flight Wi-Fi influences my choice of airline	1	2	3	4	5

*In this survey, an “in-flight meal” refers to a simple, pre-packaged meal or snack offered during the flight (e.g., a breakfast deal with a sandwich, fruit and a drink). It does not include luxury or multi-course meals.

Thank you for your participation!

Figure B.2. Page 2 of the passenger survey.

Appendix C

Data Analysis Workflow Pseudocode

The following pseudocode summarises the main data processing and modelling workflow used in this research, starting from the point at which survey responses were collected and cleaned. This pseudocode is designed to show the sequence and logic of operations, without detailing the specific syntax of the Python code used.

Purpose:

To process survey responses, prepare data, model passenger WTP for ancillary services, evaluate model performance, and interpret feature importance.

Functional Requirements

- Inputs
 - Cleaned survey dataset (.xlsx) with demographic and travel behaviour data
- Processes
 - Encode categorical variables
 - Create WTP score and categories (Low, Medium, High)
 - Split into training/testing data
 - Train ML models (Logistic Regression, Random Forest, SVM, XGBoost)
 - Evaluate model performance
 - Generate confusion matrices for best models
 - Perform SHAP & permutation importance analysis

- Outputs
 - Performance metrics table for all models and ancillaries
 - Confusion matrices
 - Feature importance plots and rankings
 - Summary of key predictors per ancillary

Main Steps (Algorithm)

```
1. Load survey data from file into dataframe

2. Encode categorical variables (Gender, Age, Annual Income,
   Usual Travel Purpose, Annual Flight Frequency) into numeric
   codes

3. For each ancillary service:
   3.1 Generate composite WTP score from survey answers
   3.2 Categorise WTP into:
       0 = Low      (1 ≤ score ≤ 2.5)
       1 = Medium  (2.5 < score ≤ 3.5)
       2 = High    (3.5 < score ≤ 5)

-----
# A. Predictive Modeling Pipeline

4. Prepare Features and Labels
   4.1 Select predictors (encoded demographic & travel
       variables) as X
   4.2 Set target (y) = selected ancillary WTP class

5. Train-test split (80% training / 20% testing, stratified)
```

```
6. If model requires scaling (Logistic Regression, SVM):
    6.1 Fit StandardScaler on X_train
    6.2 Transform X_train and X_test

7. Define ML models: Logistic Regression, Random Forest,
   SVM (RBF kernel), XGBoost

8. For each model:
    8.1 Train model on training data
    8.2 Predict on test data
    8.3 Calculate performance metrics: Accuracy, Precision,
        Recall, F1 Score, ROC-AUC
    8.4 Store metrics in results table

9. Identify best model for ancillary (highest F1 score)

10. Error Analysis:
    10.1 Predict with best model
    10.2 Generate confusion matrix (Low/Medium/High classes)
    10.3 Visualise matrix as heatmap

-----
# B. Feature Importance Analysis

11. Feature Importance Analysis (per ancillary):
    11.1 Train XGBoost regressor/classifier on full dataset
    11.2 Compute SHAP values, plot summary
    11.3 Compute permutation importance, rank features
    11.4 Save plots and rankings
```

Workflow Diagram

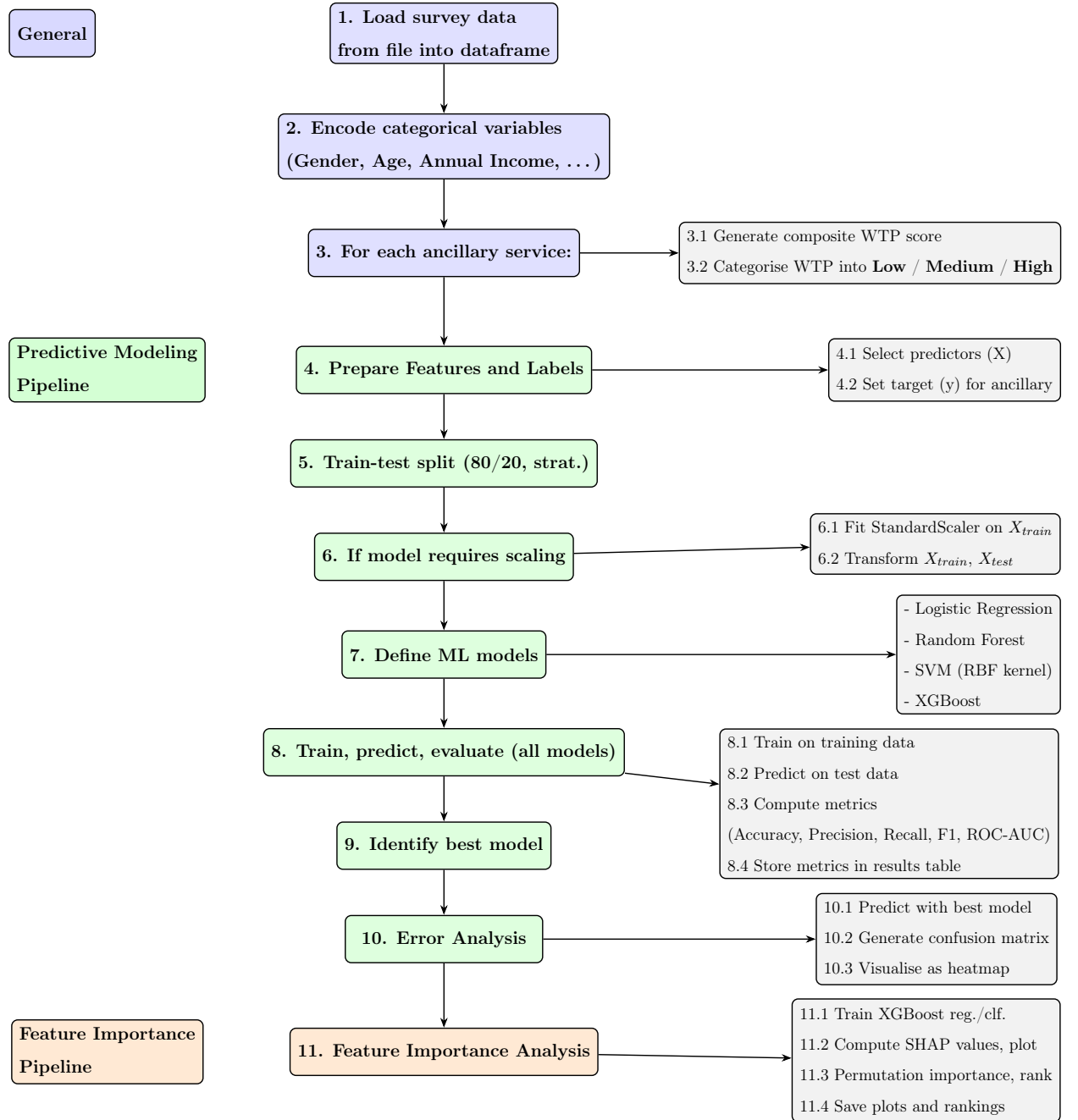


Figure C.1. Visual flowchart of the data analysis workflow.