



Table of Contents

[Background & Question 2](#_gjdgxs)

[Question: 2](#_30j0zll)

[Hypothesis and Prediction 2](#_3dy6vkm)

[Hypothesis: 2](#_1t3h5sf)

[Prediction: 2](#_4d34og8)

[Methods 2](#_2s8eyo1)

[Outcome Variable: 2](#_26in1rg)

[Predictor Variables: 2](#_lnxbz9)

[Results 5](#_z337ya)

[Overview of Variables: 5](#_3j2qqm3)

[Descriptive Statistics: 5](#_1y810tw)

[Demographic Summary Table: 5](#_rw3pmqgv7voq)

[Demographic Summary Table by Country: 6](#_rmafa2qjz8mk)

[AI Use and Satisfaction Summary: 7](#_ihnq7srjso3e)

[AI Use and Satisfaction Summary by Country: 8](#_6uz37y4f4is7)

[Visualizations: 8](#_4i7ojhp)

[Frequency Bar Charts of Demographic Variables: 8](#_3571t6n0okg8)

[Pie Charts of AI Tool Usage: 9](#_gd3zms6qlkx6)

[Stacked Bar Graphs by AI Satisfaction: 9](#_mlv9okwf6x8o)

[Scatterplot of AI Trust vs. AI Satisfaction: 10](#_9o5rb2oajbis)

[Histogram of AI Usage Levels: 10](#_aj7hcwa8bh5t)

[**Discussion and Next Steps 10**](#_6cjx1tk0ob6n)

[Summary of Key Takeaways: 10](#_b1bqdpl5ukl0)

[Data Cleaning and Preprocessing Plan: 10](#_eu5jkuuxy2rz)

[References 11](#_2xcytpi)

[**Appendix 12**](#_m3adjyvj41xm)

# Background & Question

## Question:

Can we predict customer satisfaction with AI content to help businesses identify the type of customer base most likely to have a positive experience with AI tools they deploy?

## Hypothesis and Prediction

### Hypothesis:

While the implementation of AI in retail settings improves cost efficiency and reduces errors (Kasareni, 2021), excessive reliance on AI in customer service without human interaction will reduce customer satisfaction.

### Prediction:

Customer satisfaction with AI-driven customer service will be higher among specific demographic groups and for certain AI tools.

# Methods

Dataset Description and Acquisition:

The “Customer Satisfaction Response to Artificial Intelligence Tools Usage During Online Shopping” dataset has much of the data that we will need to analyze customer satisfaction and use of AI. Our dataset includes customer demographic and behavioral variables from three different countries, focusing on interactions and satisfaction with AI tools. It includes 23 columns and 656 rows, covering variables such as country, age, gender, education, preferred online services, annual salary, payment methods, AI endorsement and privacy perceptions, and specific AI tools used.

<https://figshare.com/articles/dataset/Customer_Satisfaction_Response_to_Artificial_Intelligence_Tools_Usage_During_Online_Shopping/24633105?file=43284342>

### Outcome Variable:

Customer Satisfaction (AI\_Satisfaction)

### Predictor Variables:

1. Demographic: Country, Generation (Age), Gender, Education, and Living\_Region
2. AI tools used: Chatbots, Virtual Assistants, Voice & Photo Search
3. AI\_Trust: (AI\_Privacy\_Trust + AI\_Endorsement) / 2
4. AI\_Enhance\_Experience
5. AI\_Usage: AI tools (recode for high, moderate, or low)

Data Cleaning and Preparation:

To prepare the dataset for analysis, we cleaned and prepared the dataset for analysis by handling missing values, encoding our categorical variables, normalizing our data, inverting AI\_Privacy\_No\_Trust to become AI\_Privacy\_Trust and creating new variables such as AI\_Trust which will combine AI\_Privacy\_Trust and AI\_Endorsement and AI\_Usage which allows us to see exposure to different variations of AI tools (high, moderate, and low exposure). We also grouped AI use by country since some countries can overall have differing opinions about AI than others. We will also remove records of individuals who reported they do not shop online, as well as the single data point missing a gender report. This ensures our analysis remains focused on the online sales context where AI tools are used. This data only constitutes 3% of the overall data, so this cleaning step is unlikely to impact the overall results and also ensures a relevant sample to provide to online retailers. .

Exploratory Data Analysis Approach:

For EDA, we will use a combination of summary statistics, visualizations, and correlation analysis to explore relationships within the data. We will be using classification based machine learning models to look at when people are satisfied with using AI and when they are dissatisfied. Some of the models we will use are Logistic Regression, Decision Trees, Random Forests, SVM, KNN, and Neural Networks. The “AI\_Satisfication” column will be our predictor. We will use threshold analysis to identify the point when AI usage begins to correlate with reduced satisfaction and examine feature importance analysis to identify which demographics and AI tools are most associated with high satisfaction. We will also analyze the model for high predictive accuracy, precision, recall, and AUC.

Data Dictionary:

| Variable | Data Type | Description |
| --- | --- | --- |
| Country | String | Country of the respondent |
| Online\_Consumer | String | Whether the consumer shops online (YES/NO) |
| Age | String | Age group of the respondent (Gen X, Gen Z, etc.) |
| Annual\_Salary | String | Salary range of the respondent (Low, Medium, High) |
| Gender | String | Gender of the respondent (Male/Female) |
| Education | String | Highest education level achieved by the respondent (Masters' Degree, University Graduate, etc.) |
| Payment\_Method\_Credit/Debit | String | Whether the respondent uses credit or debit cards for payments (YES/NO) |
| Living\_Region | String | Type of living area (Metropolitan, Rural Areas) |
| Online\_Service\_Preference | String | Preference for using online services (YES/NO) |
| AI\_Endorsement | String | Whether the respondent endorses the use of AI in online shopping (YES/NO) |
| AI\_Privacy\_Trust  (Inversion of AI\_Privacy\_No\_Trust) | String | Whether the respondent has privacy concerns or lacks trust in AI (YES/NO)  Will take the customer's responses and flip them (If a yes, will now become a no and vice versa) |
| AI\_Enhance\_Experience | String | Whether the respondent feels AI enhances their shopping experience (YES/NO) |
| AI\_Satisfaction | String | Overall satisfaction level |
| AI\_Trust | String | (AI\_Privacy\_No\_Trust + 1 + AI\_Endorsement+1)/2 (Will create High (2), Medium (1), Low Trust(0)) |
| AI\_Usage | String | AI\_Tools variables/3 (Will create High (3), Medium (2), Low Usage(1)) |
| AI\_Tools\_Used\_Chatbots | String | Use of chatbots for assistance (YES/NO) |
| AI\_Tools\_Used\_Virtual\_Assistant | String | Use of virtual assistants (YES/NO) |
| AI\_Tools\_Used\_Voice&Photo\_Search | String | Use of voice and photo search tools (YES/NO) |
| Paymeny\_Method\_COD | String | Use of Cash on Delivery as a payment method (YES/NO) |
| Payment\_Method\_Ewallet | String | Use of e-wallets as a payment method (YES/NO) |
| Product\_Category\_Appliances | String | Interest or purchase in appliances category (YES/NO) |
| Product\_Category\_Electronics | String | Interest or purchase in electronics category (YES/NO) |
| Product\_Category\_Groceries | String | Interest or purchase in groceries category (YES/NO) |
| Product\_Category\_Personal\_Care | String | Interest or purchase in personal care category (YES/NO) |
| Product\_Category\_Clothing | String | Interest or purchase in clothing category (YES/NO) |

# Results

## Overview of Variables:

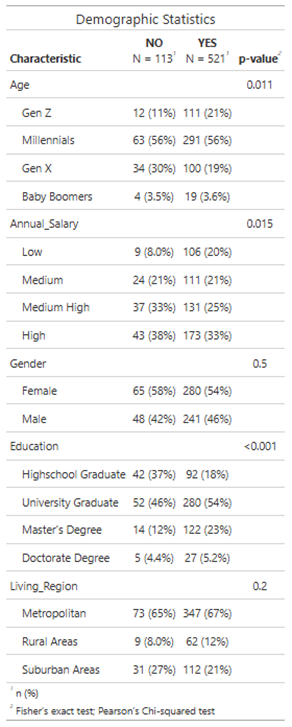
All of our variables are categorical so we will summarize them using frequencies, percentages, and proportions in tables and graphs.

## Descriptive Statistics:

### 

### Demographic Summary Table:

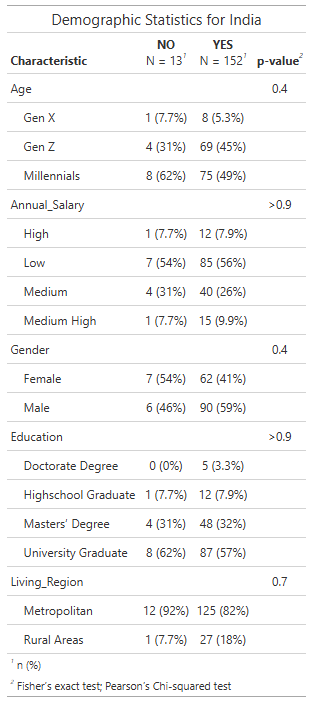
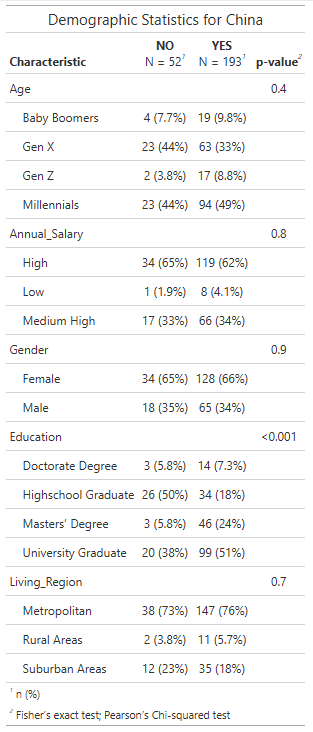
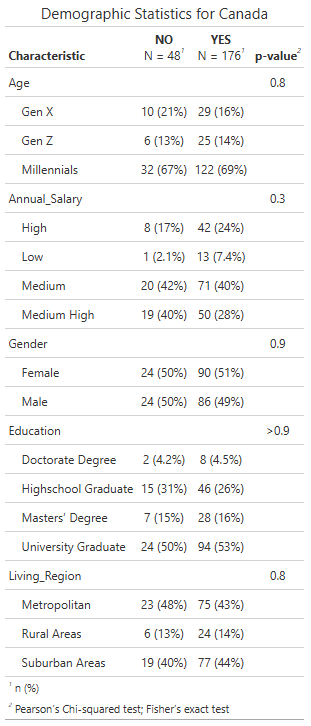
The demographic statistics for the full dataset show significant differences in AI satisfaction across age, annual salary, and education levels. Gen Z, individuals with low annual salaries, and those with higher levels of education, specifically Master’s degrees report higher satisfaction with AI. There were no significant differences in AI satisfaction based on gender or living region.



*Figure 1 Demographic Summary Table*

### Demographic Summary Table by Country:

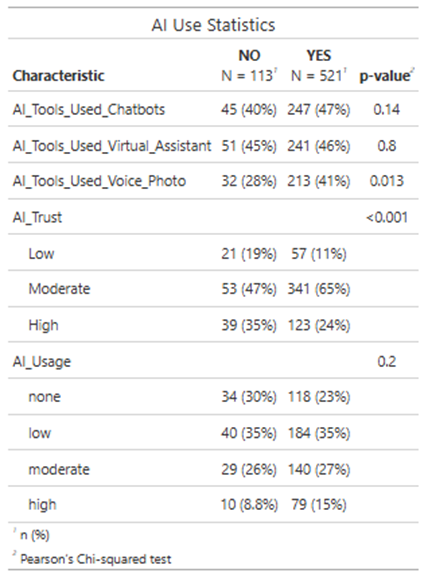
The Demographic Statistics tables for Canada, China, and India reveal country-specific differences in AI satisfaction, with China showing statistically significant differences across Education but other demographic variables showing no statistically significant differences in AI Satisfaction.



*Figure 2 Demographic Summary Table by Country*

### AI Use and Satisfaction Summary:

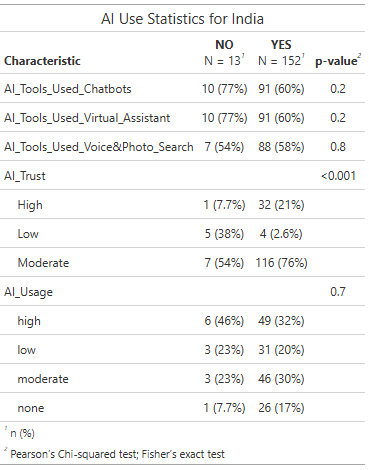
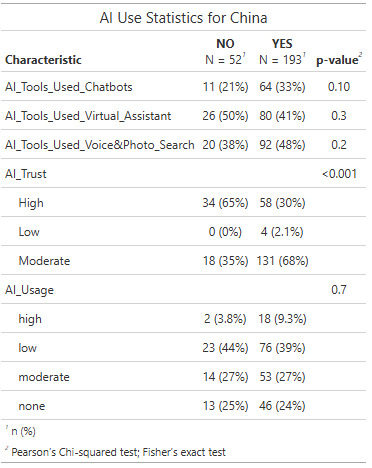
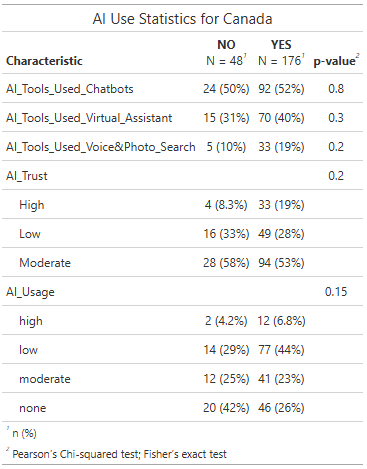
This table illustrates overall AI usage trends from the full dataset which indicates a statistically significant relationship with the use of Voice and Photo Search AI tools.



*Figure 3 AI Use and Satisfaction Summary*

### AI Use and Satisfaction Summary by Country:

The AI Use Statistics tables for Canada, China, and India reveal country-specific differences in AI adoption, with China and India showing statistically significant levels of AI trust, while Canada does not.



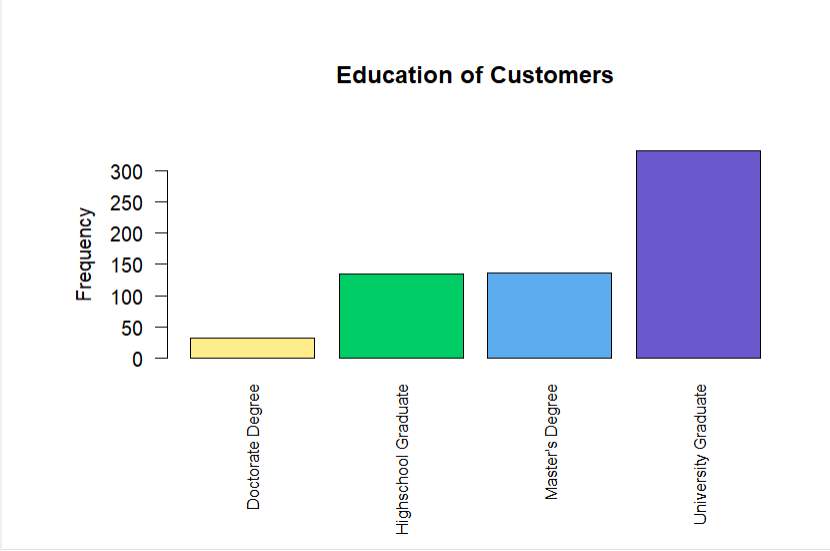
*Figure 4 AI Use and Satisfaction Summary by Country*

## Visualizations:

### Frequency Bar Charts of Demographic Variables:

This table will present frequencies and percentages of key demographic variables, such as country, age, gender, education, and online service preferences.



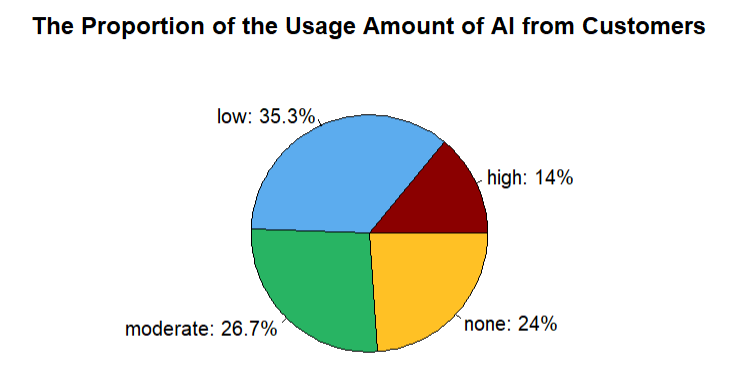


*Figure 5 Frequency Bar Charts - Demographics*

Interpretation: From these demographics we can get a better understanding of our dataset. From our data most of our customers come from China and Canada, with a few stemming from India. Our sample mostly prefers online shopping to shopping at a store, and are in the generation of millennials (born 1981 - 1996), and have mostly graduated college or university with a bachelor's degree.. Our sample is also pretty evenly split in terms of gender with it being roughly 55% female and 45% male. Finally most of the customers in our dataset have either a medium high or high annual salary.

### Pie Charts of AI Tool Usage:

These pie charts will show the proportion of customers in high, moderate, and low AI usage categories.

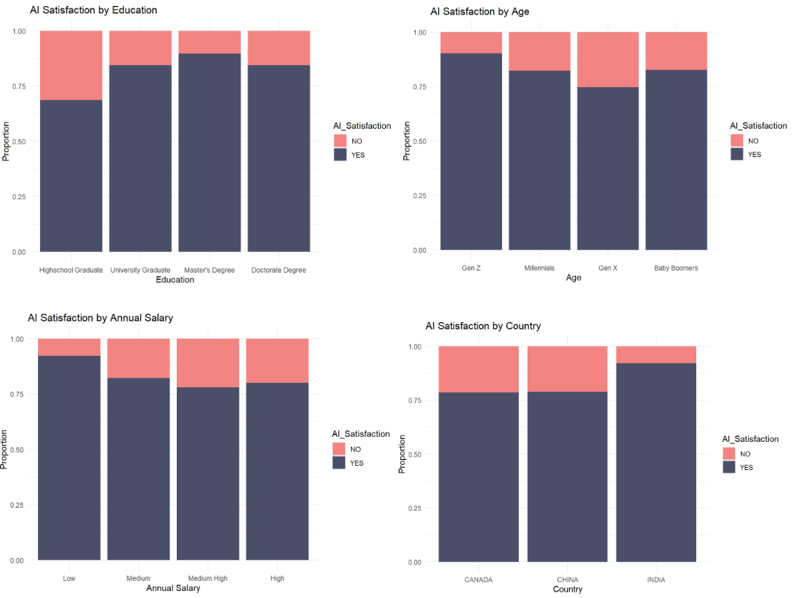


*Figure 6 Pie Chart - AI Tool Usage*

Interpretation: From this graph we can see that 35% use a low amount of AI, 27% use a moderate amount of AI, 14% use a high amount of AI, and finally 24% do not use artificial intelligence.

### Stacked Bar Graphs by AI Satisfaction:

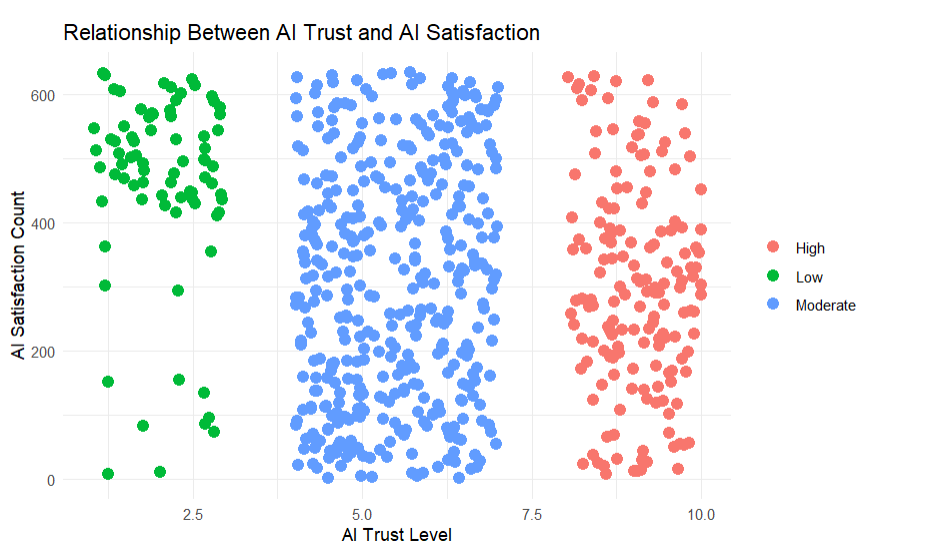
The stacked bar graphs by variable (Education, Annual Salary, Age, and Country) indicate that the highest proportion of reported AI Satisfaction are individuals with a Master’s degree, while those with only a high school education have the lowest. In terms of salary, those with a medium-high salary have the lowest satisfaction with AI and those with a low salary report the highest satisfaction. Gen Z shows the highest satisfaction with AI, while Gen X has the lowest and India has a higher proportion of "yes" responses for AI satisfaction compared to Canada or China.



*Figure 7 Stacked Bar Graphs AI Satisfaction*

### Scatterplot of AI Trust vs. AI Satisfaction:

This scatter plot will visualize the relationship between the newly created AI\_Trust variable and AI satisfaction levels.

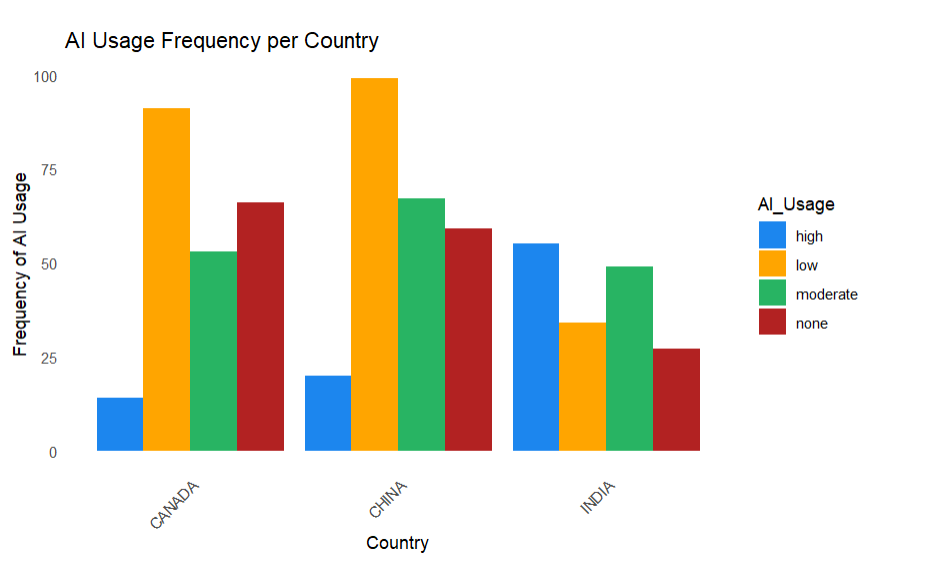


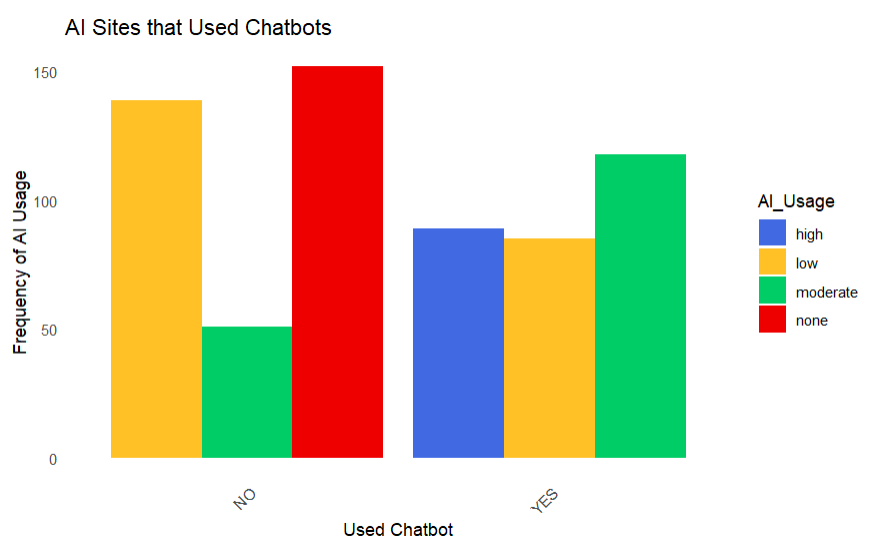
*Figure 8 Scatterplot of AI Trust vs. AI Satisfaction*

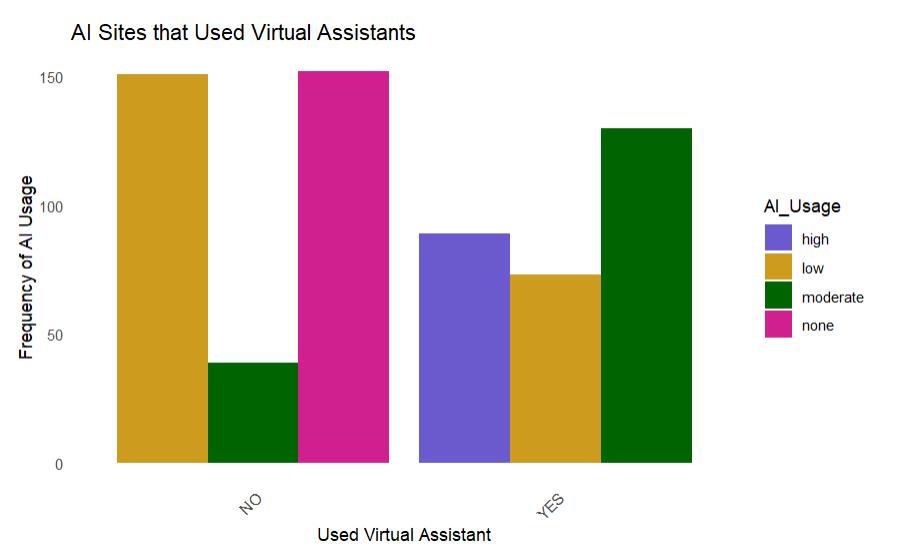
Interpretation: While there is no distinct trend between AI trust and AI satisfaction we can see that generally the more trust a person has then the more satisfied they are. For this graph trust is listed on a scale of 1 - 10 and people with 4 or above usually have a moderate or high amount of trust with their AI usage.

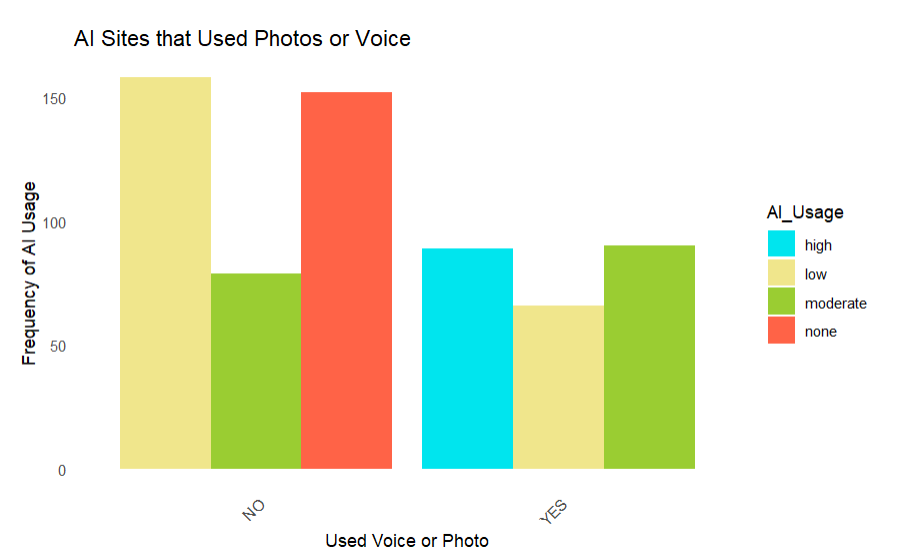
### Histogram of AI Usage Levels:

This histogram will show the distribution of AI usage levels (high, moderate, low) across the dataset.





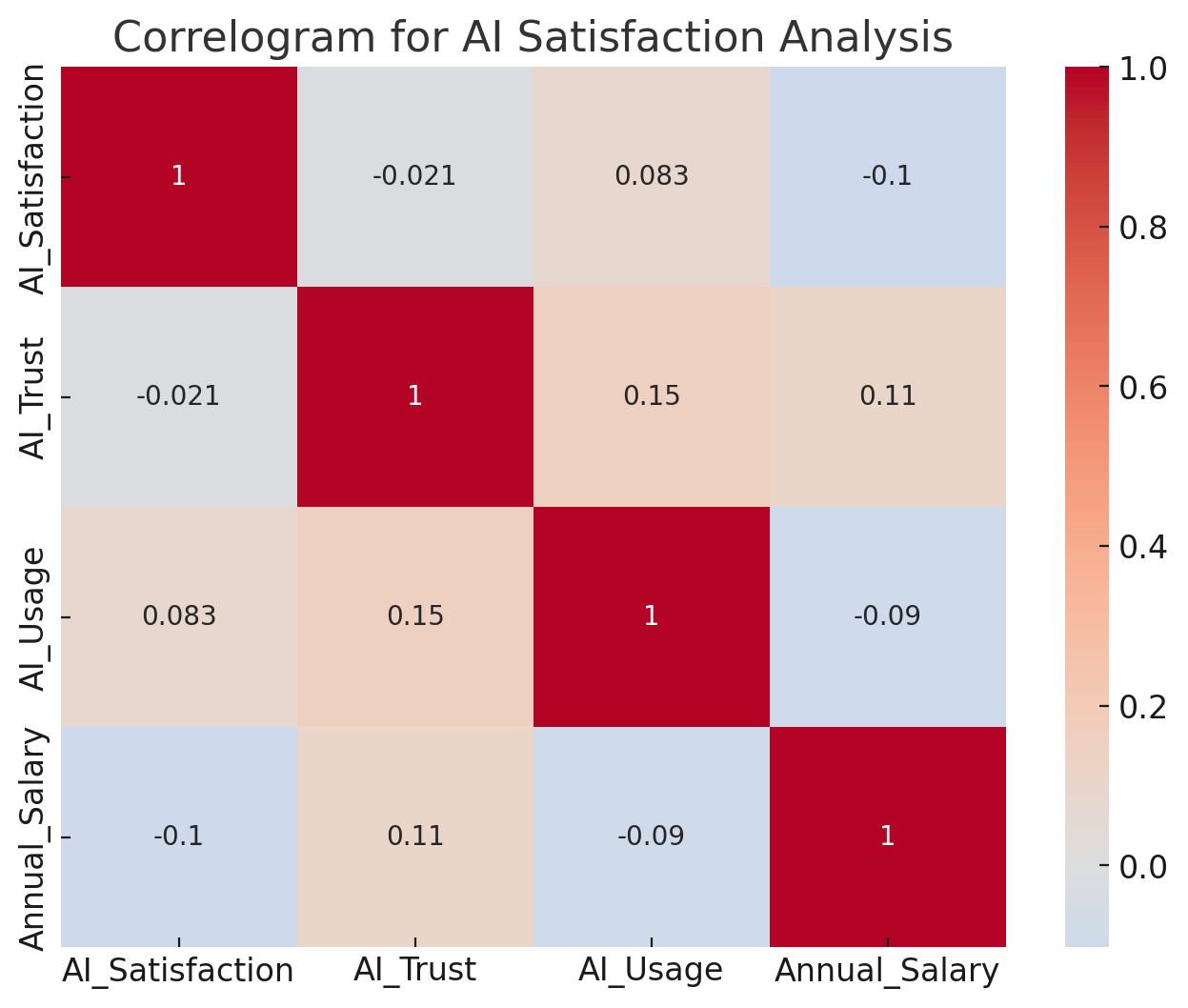




*Figure 9 Histogram of AI Usage Levels*

Interpretation: Within the three countries our data set has, both China and Canada mostly use a low amount of AI or no AI. India on the other hand uses high or moderate amounts of AI. This is interesting because India has less sampled customers but they still are using the most AI. The next three graphs show what AI is being used and how often it is used. Based on the results we can see that a moderate amount of chatbots and virtual assistants are being used, and a high amount of AI that uses voice or photo is being used. This could be based on the fact that AIs with voice or photo recognition have been around longer and are used everyday within our smartphones, while chatbots with AI have been a more recent invention.

Multicollinearity:



*Figure 10 AI Satisfaction Correlogram*

Interpretation: The correlogram shows that there is not a high correlation between AI satisfaction and some of the other key variables. This is key because you do not want your predictors to have a high correlation with your target variable.

# Discussion and Next Steps

## Summary of Key Takeaways:

Based on the question, "Can we predict customer satisfaction with AI content to help businesses identify the type of customer base most likely to have a positive experience with AI tools they deploy?" Our findings indicate that demographic factors and specific AI tools significantly influence customer satisfaction levels. In the stacked bar graphs there is clear evidence showing an increasing trend for AI\_Satisfaction based on the education level of the customer. For those clients that have a higher education level, there is a tendency for the customer to be more comfortable with AI (Figure 7). Furthermore, Figure 2’s demographic breakdown by country highlights differences across Canada, China, and India, with notable satisfaction levels in China and distinct patterns in AI satisfaction by education levels. These findings suggest that generational factors and educational attainment may shape positive perceptions of AI in customer service, underscoring a demographic-specific response to AI that can help businesses better target their implementations.

In terms of AI tool usage, Figure 3 reveals a statistically significant positive relationship between satisfaction and the use of Voice and Photo Search tools. Country-specific analysis (Figure 4) further indicates that AI trust levels are highest in China, whereas Canada and India show more skepticism. In Figure 8’s Scatterplot of AI Trust vs. Satisfaction it would be useful to see this data broken down at the country level. This reflects a trend where greater trust in AI correlates with increased satisfaction, although the correlation is moderate. Additionally, the Histogram of AI Usage Levels (Figure 9) points to notable country-specific usage patterns, with India having the highest engagement with AI tools, primarily those that offer voice and photo functionalities. These insights suggest that satisfaction with AI tools is heavily influenced by the type of tool and the users' cultural context, which can inform businesses aiming to optimize customer experiences by targeting demographic and regional preferences.

## Data Cleaning and Preprocessing Plan:

Our next steps involve refining our data cleaning and preprocessing approach to ensure that the dataset is optimized for generating actionable insights. We’ll start by addressing any remaining inconsistencies, such as handling outliers and re-checking for missing values that may impact our analysis. For categorical variables, we’ll streamline encoding, ensuring consistency across variables like age groups, education levels, and salary categories to facilitate smoother model interpretations and comparisons. Additionally, we will further optimize our AI\_Trust and AI\_Usage variables to allow us to better capture nuanced relationships, particularly by adjusting these variables for any skewness and exploring interactions between trust and satisfaction at more granular levels.

To enhance interpretability and visualization, we plan to improve the graphical representation of key findings to deliver clearer insights. Our stacked bar graphs, histograms, and scatter plots will be redesigned for higher visual appeal and readability, using distinct color schemes and sharper labels. For instance, we’ll refine Figure 5’s demographic bar charts by adding percentages directly to bars for clearer interpretation at a glance. Similarly, Figure 8’s scatterplot of AI Trust vs. Satisfaction could benefit from fitted lines or bands to highlight patterns more effectively. These visual upgrades, combined with the optimized data, should enhance both the aesthetic quality and the actionable clarity of insights, enabling stakeholders to more easily identify trends in AI satisfaction by demographic, tool type, and trust level

# References

Bergen, Mark, and Lynn Doan. “Tech Giants Are Set to Spend $200 Billion This Year Chasing AI.” *Bloomberg.com*, Bloomberg, Nov. 2024, www.bloomberg.com/news/articles/2024-11-01/tech-giants-are-set-to-spend-200-billion-this-year-chasing-ai. Accessed 3 Nov. 2024.

Kannan, Rathimala; Ramakrishnan, Kannan; Ersoy, Ayse Begum; Contu, Davide (2023). Customer Satisfaction Response to Artificial Intelligence Tools Usage During Online Shopping. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.24633105.v1>

Kasaraneni, Ramana Kumar (2021). AI-Enhanced Supply Chain Collaboration Platforms for Retail: Improving Coordination and Reducing Costs. *Journal of Bioinformatics and Artificial Intelligence*, 1(1), pp. 410–450. Available at: https://biotechjournal.org/index.php/jbai/article/view/98 (Accessed: 3 November 2024).

# Appendix

DSE\_Capstone\_EDA\_Canfield

Ryan Canfield

2024-11-10

## **R Markdown**

library(tidyverse) # For data manipulation.

library(ggplot2) # For data visualizations.  
library(RColorBrewer) # For coloring visuals.

# Reading the CSV file into a data frame  
 df <- read.csv("cleaned\_data.csv")  
  
 # Display a preview of the data frame.  
 head(df)

# **Visualizations:**

## **Frequency Bar Charts of Demographic Variables:**

### **This table will present frequencies and percentages of key demographic variables, such as country, age, gender, education, and online service preferences.**

# Specify colors for each level  
 country\_colors <- c("red", "gold", "orange")  
 age\_colors <- c("navy", "darkorchid4", "deeppink3", "salmon")  
 gender\_colors <- c("pink", "skyblue")   
 education\_colors <- c("lightgoldenrod1", "springgreen3", "steelblue2", "slateblue3")   
 OnlineS\_colors <- c("darkred", "lightgreen")   
 salary\_colors <- c("deepskyblue", "darkolivegreen", "hotpink4", "firebrick3")  
  
  
  
 par(mfrow=c(2,3))  
  
 # Plot 1: Country  
 barplot(table(df$Country),  
 col = country\_colors,   
 xlab = "Country",  
 ylab = "Frequency",  
 main = "Customer Origins",  
 cex.names = 0.8)  
  
 # Plot 2: Online Service Preference  
 barplot(table(df$Online\_Service\_Preference),  
 col = OnlineS\_colors,  
 xlab = "Online Shopping",  
 ylab = "Frequency",  
 main = "Do They Prefer Online Shopping",  
 cex.names = 0.8)  
  
 # Plot 3: Gender  
 barplot(table(df$Gender),  
 col = gender\_colors,  
 xlab = "Gender",  
 ylab = "Frequency",  
 main = "Gender of Customer",  
 cex.names = 0.8)  
  
 # Plot 4: Age  
 barplot(table(df$Age),  
 col = age\_colors,  
 xlab = "Age",  
 ylab = "Frequency",  
 main = "Age of Customers",  
 las = 2,   
 cex.names = 0.8)  
   
  
 # Plot 5: Annual Salary  
 barplot(table(df$Annual\_Salary),  
 col = salary\_colors,  
 xlab = "Salary",  
 ylab = "Frequency",  
 main = "Customers Salary",  
 las = 2,  
 cex.names = 0.8)

par(mar=c(5, 4, 4, 2) + 2.4) # Adjusts the bottom margin

# Plot 6: Education  
 barplot(table(df$Education),  
 col = education\_colors,  
 ylab = "Frequency",  
 main = "Education of Customers",  
 las = 2,   
 cex.names = 0.8)

## **Pie Charts of AI Tool Usage:**

### **These pie charts will show the proportion of customers in high, moderate, and low AI usage categories.**

# Calculate the percentage of people who use AI  
 AI\_usage\_percents <- df %>%  
 count(AI\_Usage) %>%  
 mutate(Percentage = n / sum(n) \* 100)  
  
 # Create the pie chart using base R  
 pie(AI\_usage\_percents$Percentage,  
 labels = paste0(AI\_usage\_percents$AI\_Usage, ": ", round(AI\_usage\_percents$Percentage, 1), "%"),  
 col = c('darkred', 'steelblue2', '#28B463', 'goldenrod1'), # Customize colors  
 main = "The Proportion of the Usage Amount of AI from Customers")

## **Scatterplot of AI Trust vs. AI Satisfaction:**

### **This scatter plot will visualize the relationship between the newly created AI\_Trust variable and AI satisfaction levels.**

unique(df$AI\_Trust)

## [1] "Moderate" "Low" "High"

unique(df$AI\_Satisfaction)

## [1] "YES" "NO"

set.seed(3) # For repeated results  
  
 # Function to generate random number based on category  
 generate\_random <- function(category) {  
 if (category == "Low") {  
 return(round(runif(1, min = 1, max = 3), 2)) # Random number between 1 and 3  
 } else if (category == "Moderate") {  
 return(round(runif(1, min = 4, max = 7), 2)) # Random number between 4 and 7  
 } else if (category == "High") {  
 return(round(runif(1, min = 8, max = 10), 2)) # Random number between 8 and 10  
 }  
 }  
  
  
 df$Random\_Num\_based\_AITrust <- sapply(df$AI\_Trust, generate\_random)  
  
 # View the result  
 head(df)

df$Index <- seq\_along(df$AI\_Trust)  
  
  
 ggplot(df, aes(x = Random\_Num\_based\_AITrust, y = Index)) +  
 geom\_point(aes(color = AI\_Trust), size = 3) +  
 labs(  
 title = "Relationship Between AI Trust and AI Satisfaction",  
 x = "AI Trust Level",  
 y = "AI Satisfaction Count"  
 ) +  
 theme\_minimal() +  
 theme(legend.title = element\_blank())

## **Histogram of AI Usage Levels:**

### **This histogram will show the distribution of AI usage levels (high, moderate, low) across the dataset.**

# Summarize the data to get frequencies  
 data\_summary <- df %>%  
 count(Country, AI\_Usage)  
  
 # Create the bar graph  
 ggplot(data\_summary, aes(x = Country, y = n, fill = AI\_Usage)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 labs(x = "Country", y = "Frequency of AI Usage", title = "AI Usage Frequency per Country") +  
 scale\_fill\_manual(values = c("high" = "dodgerblue2", "moderate" = "#28B463", "low" = "orange", "none" = "firebrick")) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1),  
 panel.grid = element\_blank())

# Summarize the data to get frequencies  
 data\_summary <- df %>%  
 count(AI\_Tools\_Used\_Chatbots, AI\_Usage)  
  
 # Create the bar graph  
 ggplot(data\_summary, aes(x = AI\_Tools\_Used\_Chatbots, y = n, fill = AI\_Usage)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 labs(x = "Used Chatbot", y = "Frequency of AI Usage", title = "AI Sites that Used Chatbots") +  
 scale\_fill\_manual(values = c("high" = "royalblue", "moderate" = "springgreen3", "low" = "goldenrod1", "none" = "red2")) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1),  
 panel.grid = element\_blank())

# Summarize the data to get frequencies  
 data\_summary <- df %>%  
 count(AI\_Tools\_Used\_Virtual\_Assistant, AI\_Usage)  
  
 # Create the bar graph  
 ggplot(data\_summary, aes(x = AI\_Tools\_Used\_Virtual\_Assistant, y = n, fill = AI\_Usage)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 labs(x = "Used Virtual Assistant", y = "Frequency of AI Usage", title = "AI Sites that Used Virtual Assistants") +  
 scale\_fill\_manual(values = c("high" = "slateblue", "moderate" = "darkgreen", "low" = "goldenrod3", "none" = "violetred")) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1),  
 panel.grid = element\_blank())

# Summarize the data to get frequencies  
 data\_summary <- df %>%  
 count(AI\_Tools\_Used\_Voice\_Photo, AI\_Usage)  
  
 # Create the bar graph  
 ggplot(data\_summary, aes(x = AI\_Tools\_Used\_Voice\_Photo, y = n, fill = AI\_Usage)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 labs(x = "Used Voice or Photo", y = "Frequency of AI Usage", title = "AI Sites that Used Photos or Voice") +  
 scale\_fill\_manual(values = c("high" = "turquoise2", "moderate" = "yellowgreen", "low" = "khaki", "none" = "tomato")) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1),  
 panel.grid = element\_blank())

## **Correlogram:**

**Will show the correlation between target variable and other variables**

unique\_values = {col: data[col].unique() for col in data.columns if col in ["AI\_Satisfaction", "AI\_Trust", "AI\_Usage", "Annual\_Salary"]}

unique\_values

import seaborn as sns

import matplotlib.pyplot as plt

# Define mappings for categorical to numerical conversion

mappings = {

'AI\_Satisfaction': {'YES': 1, 'NO': 0},

'AI\_Trust': {'Low': 1, 'Moderate': 2, 'High': 3},

'AI\_Usage': {'none': 0, 'low': 1, 'moderate': 2, 'high': 3},

'Annual\_Salary': {'Low': 1, 'Medium': 2, 'Medium High': 3, 'High': 4}

}

# Apply the mappings

for col, mapping in mappings.items():

data[col] = data[col].map(mapping)

# Select columns for the correlogram, focusing on those that are now numeric

numeric\_columns = ['AI\_Satisfaction', 'AI\_Trust', 'AI\_Usage', 'Annual\_Salary']

correlation\_data = data[numeric\_columns]

# Generate a correlogram

plt.figure(figsize=(8, 6))

sns.heatmap(correlation\_data.corr(), annot=True, cmap='coolwarm', center=0)

plt.title('Correlogram for AI Satisfaction Analysis')

plt.show()