AisleMate:

YOUR GUIDE TO AISLES AND PRODUCTS

Agenda

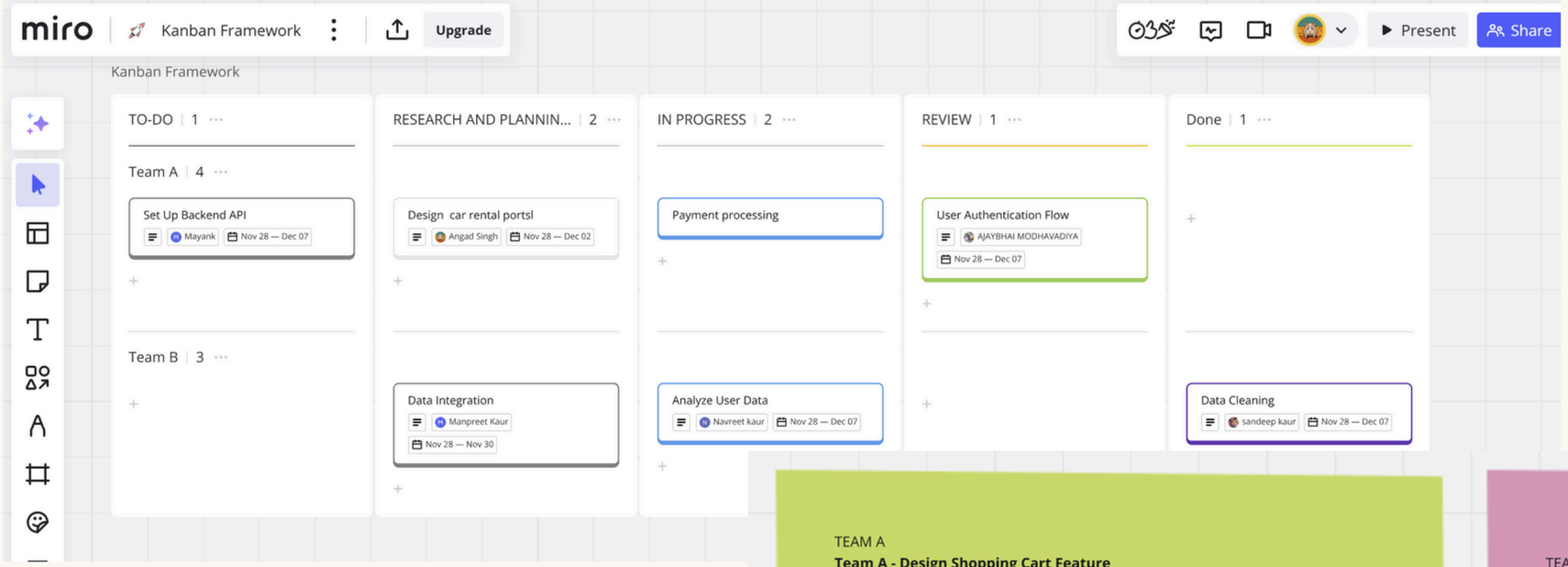
1. Project Scope
2. What is our Objective?
3. Project Planning
4. Resource Allocation
5. Data Collection
6. Data Preprocessing
7. Data Exploration
8. Data Analysis
9. Data Modeling
10. Data Visualization
11. Conclusion

BEFORE STARTING PRESENTATION HERE IS OUR KANBAN & MIRO BOARD LINKS

- PURPOSE OF THE **KANBAN BOARD LINK:**
  - PROVIDES DIRECT ACCESS TO THE KANBAN BOARD USED BY THE TEAM.
  - ENSURES TEAM MEMBERS AND STAKEHOLDERS CAN VIEW AND TRACK PROJECT PROGRESS IN REAL TIME.
- LINK TO MIRO KANBAN BOARD:
  - MIRO KANBAN BOARD LINK
- LINK TO MIRO BOARD (**REQUIREMENT ANALYSIS**)

  MIRO BOARD LINK

KANBAN & MIRO BOARD SCREENSHOT



TEAM A

Team A - Design Shopping Cart Feature
Task Title: Design Shopping Cart Feature.
Description: This task involves creating wireframes and mockups for the shopping cart page, ensuring the user interface is user-friendly and visually appealing.
Assigned To: **Angad Singh** - As the Project Manager, Angad Singh is directly handling the initial design elements for the shopping cart feature, guiding the visual aspects of the application.

Team A - Set Up Backend API
Task Title: Set Up Backend API.
Description: Develop initial API endpoints for user data retrieval, including database connections and server integration.
Assigned To: **Mayank** - As the Data Scientist, Mayank will also handle aspects of backend development related to API setup, ensuring data flows seamlessly to support features.

Team A - User Authentication Flow
Task Title: User Authentication Flow.
Description: Implement secure user login and registration processes, ensuring identity verification for app users.
Assigned To: **Ajaybhai Modhavadhiya** - Assigned to Ajaybhai, who will handle the technical implementation of secure authentication flows.

TEAM B

Team B - Analyze User Data
Task Title: Analyze User Data.
Description: Conduct data analysis to gain insights into user preferences and behaviors that can help in optimizing app features.
Assigned To: **Navreet Kaur** - Navreet, as part of the Data Engineering team, is responsible for gathering insights from the data, supporting data-driven feature development.

Team B - Data Integration
Task Title: Data Integration.
Description: Integrate data from various sources to ensure a complete dataset for analysis and feature implementation.
Assigned To: **Manpreet Kaur** - Manpreet will handle integrating data from multiple sources to create a unified backend database.

Team B - Data Cleaning
Task Title: Data Cleaning.
Description: Prepare and clean the data for further use, removing inconsistencies and ensuring accuracy.
Assigned To: **Sandeep Kaur** - Sandeep will focus on data preprocessing to support other team members in their data-related tasks.

OBJECTIVE

To develop a mobile application that helps shoppers easily navigate shopping malls, locate specific stores, and search for products within those stores, providing real-time aisle-level navigation.

PROJECT PLANNING (Team Formation)

Project Manager	Angad Singh	Oversee the project timeline, coordinate with stakeholders, and ensure goals are met.
Data Scientists:	Mayank	Analyze and model data for features like search optimization and personalized navigation.
Data Engineers:	Navreet Kaur & Ajay	Handle data collection, cleaning, integration, and backend architecture.
Business Analysts	Manpreet Kaur	Define requirements, validate outputs, and align features with business objectives.
Developers	Sandeep Kaur	Design and develop the app interface, backend services, and navigation algorithms.

Tools & Technologies We Used

Data Collection & Processing:

- Tools: Python, SQL, Excel, and Google Sheets.
- Download dataset from Kaggle.

Data Analysis & Visualization:

- Tools: Tableau, Power BI, Python (Pandas, Matplotlib) for insights and visualizations.

Data Collection & Processing:

- Tools: Python, SQL and Excel.

	A	B	C	D	E	F
1	CustomerID	Gender	Age	Annual Income (k$)	Spending Score (1-100)	
2	1	Male	19	15	39	
3	2	Male	21	15	81	
4	3	Female	20	16	6	
5	4	Female	23	16	77	
6	5	Female	31	17	40	
7	6	Female	22	17	76	
8	7	Female	35	18	6	
9	8	Female	23	18	94	
10	9	Male	64	19	3	
11	10	Female	30	19	72	
12	11	Male	67	19	14	
13	12	Female	35	19	99	
14	13	Female	58	20	15	
15	14	Female	24	20	77	
16	15	Male	37	20	13	
17	16	Male	22	20	79	
18	17	Female	35	21	35	
19	18	Male	20	21	66	
20	19	Male	52	23	29	
21	20	Female	35	23	98	
22	21	Male	35	24	35	
23	22	Male	25	24	73	
24	23	Female	46	25	5	
25	24	Male	31	25	73	
26	25	Female	54	28	14	
27	26	Male	29	28	82	
28	27	Female	45	28	32	
29	28	Male	35	28	61	
30	29	Female	40	29	31	

OUR DATASET

Download from Kaggle

DATA CLEANING

OUTPUT

```
(CustomerID                0
 Gender                    0
 Age                       0
 Annual Income (k$)        0
 Spending Score (1-100)    0
 dtype: int64,
 0,
        CustomerID         Age  Annual Income (k$)  Spending Score (1-100)
 count  200.000000  200.000000          200.000000              200.000000
 mean   100.500000   38.850000           60.560000               50.200000
 std     57.879185   13.969007           26.264721               25.823522
 min      1.000000   18.000000           15.000000                1.000000
 25%     50.750000   28.750000           41.500000               34.750000
 50%    100.500000   36.000000           61.500000               50.000000
 75%    150.250000   49.000000           78.000000               73.000000
 max    200.000000   70.000000          137.000000               99.000000)
```

Missing Values: There are no missing values in the dataset.
Duplicates: There are no duplicate records in the dataset.
The summary statistics show that:

- The Age ranges from 18 to 70, with a mean of 38.85.
- Annual Income (k$) ranges from 15 to 137, with a mean of 60.56.
- Spending Score (1-100) ranges from 1 to 99, with a mean of 50.2.

```
(Age                        0
 Annual Income (k$)         2
 Spending Score (1-100)     0
 dtype: int64,
    CustomerID  Gender       Age  Annual Income (k$)  Spending Score (1-100)
 0           1       0  0.019231            0.000000                0.387755
 1           2       0  0.057692            0.000000                0.816327
 2           3       1  0.038462            0.008197                0.051020
 3           4       1  0.096154            0.008197                0.775510
 4           5       1  0.250000            0.016393                0.397959)
```

DATA CLEANING & PREPROCESSING

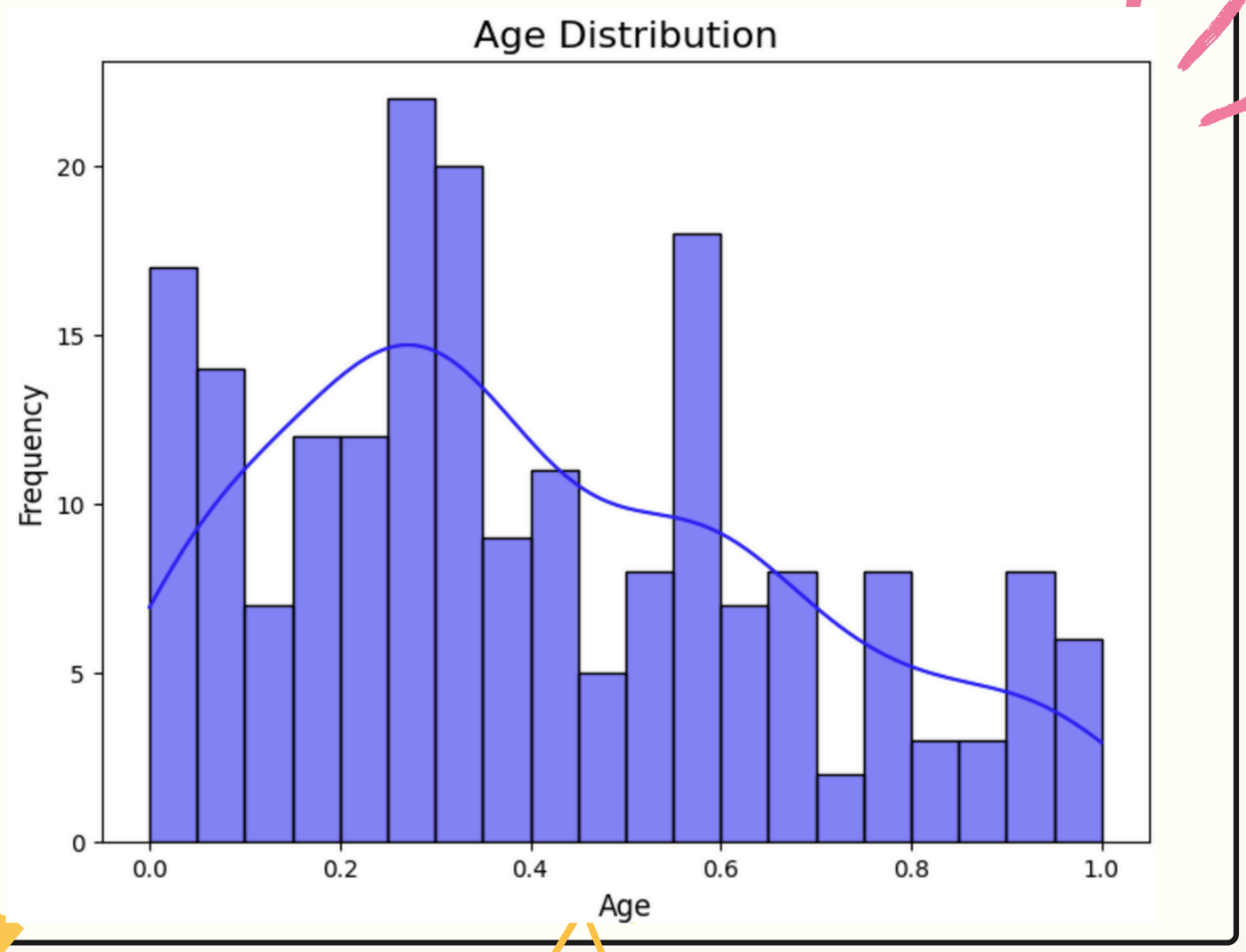
Outliers:

Annual Income (k$) has 2 outliers. No outliers were detected for Age or Spending Score (1-100).

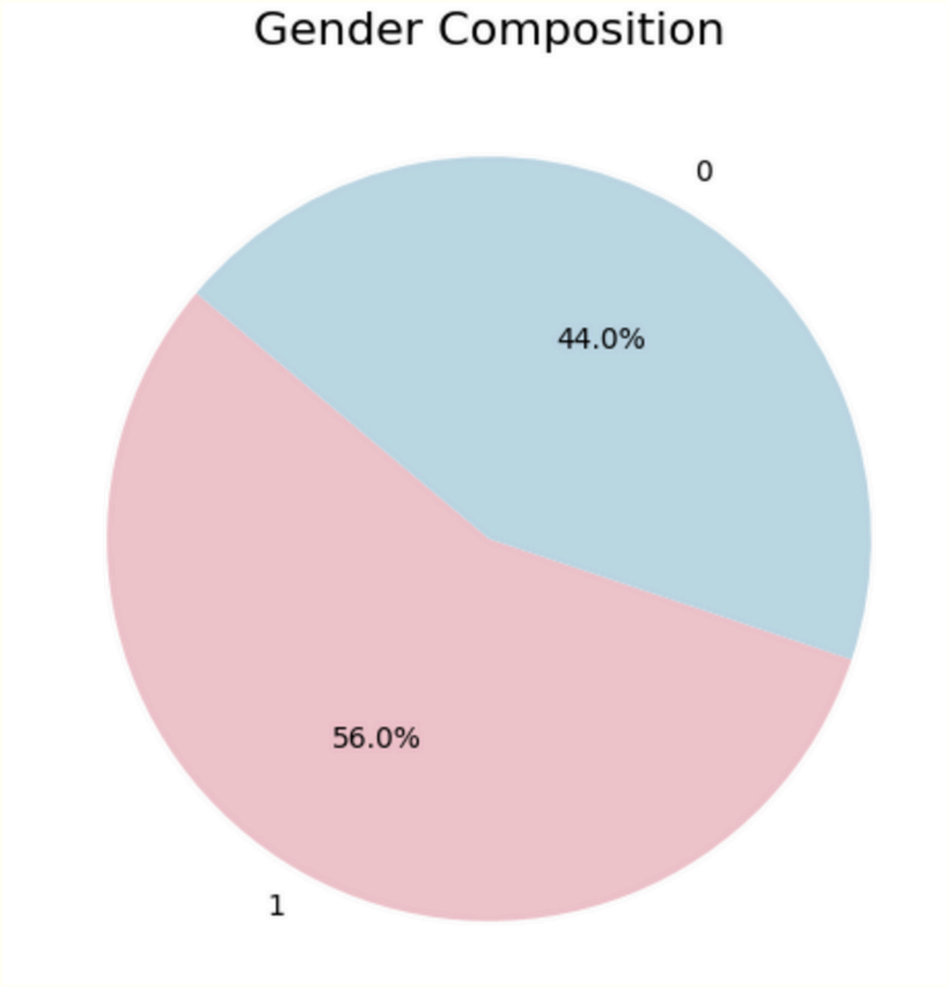
Normalization: The numerical features (Age, Annual Income, Spending Score) have been scaled between 0 and 1 using Min-Max scaling.

Encoding: The Gender column has been encoded into numerical values, where Male is mapped to 0 and Female is mapped to 1.

DATA DISTRIBUTION



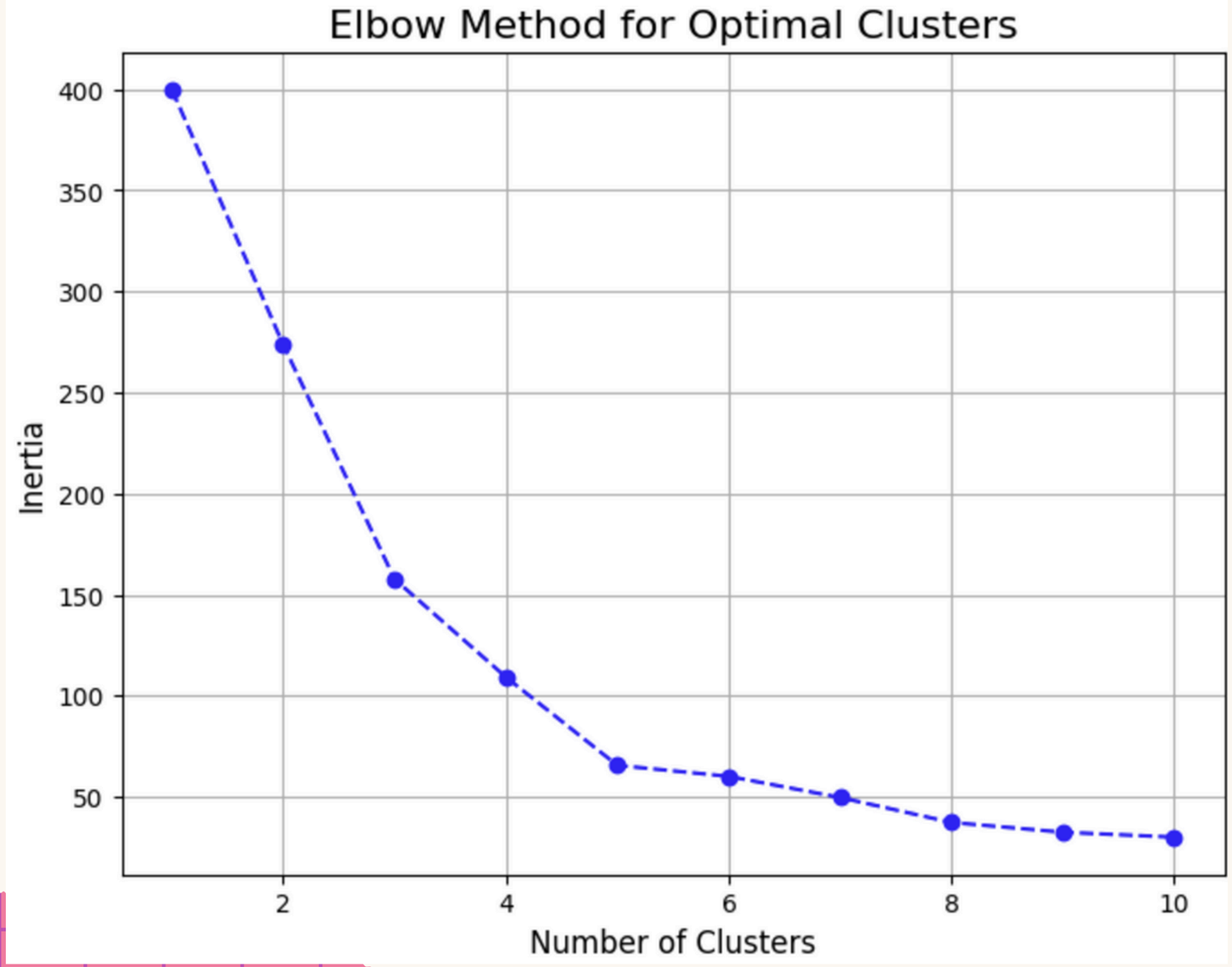
DATA DISTRIBUTION



INSIGHTS FROM VISUALIZATIONS WE GET

- Age Distribution: Most customers are between 25 and 50 years old.
- Gender Composition: The dataset is evenly split between genders.
- Annual Income vs Spending Score: A diverse relationship exists between income and spending score, with no clear linear trend. Spending behavior varies across genders and income levels.

Data Modeling



THE ELBOW METHOD PLOT INDICATES THE OPTIMAL NUMBER OF CLUSTERS, TYPICALLY AT THE "ELBOW POINT" WHERE THE RATE OF DECREASE IN INERTIA SLOWS DOWN. BASED ON THE GRAPH, A GOOD STARTING POINT MIGHT BE 4 OR 5 CLUSTERS.

Data Modeling

CUSTOMERS ARE GROUPED INTO 5 CLUSTERS BASED ON THEIR ANNUAL INCOME AND SPENDING SCORE:

- LOW INCOME & LOW SPENDING.
- HIGH INCOME & HIGH SPENDING.
- AVERAGE INCOME & VARIED SPENDING.
- HIGH SPENDING DESPITE LOW INCOME.
- MODERATE SPENDING AND INCOME.

THE SCATTER PLOT SHOWS THESE CLUSTERS VISUALLY, PROVIDING INSIGHTS INTO DIFFERENT CUSTOMER BEHAVIORS FOR STRATEGIC DECISION-MAKING.

SOME TABLEAU INSIGHTS

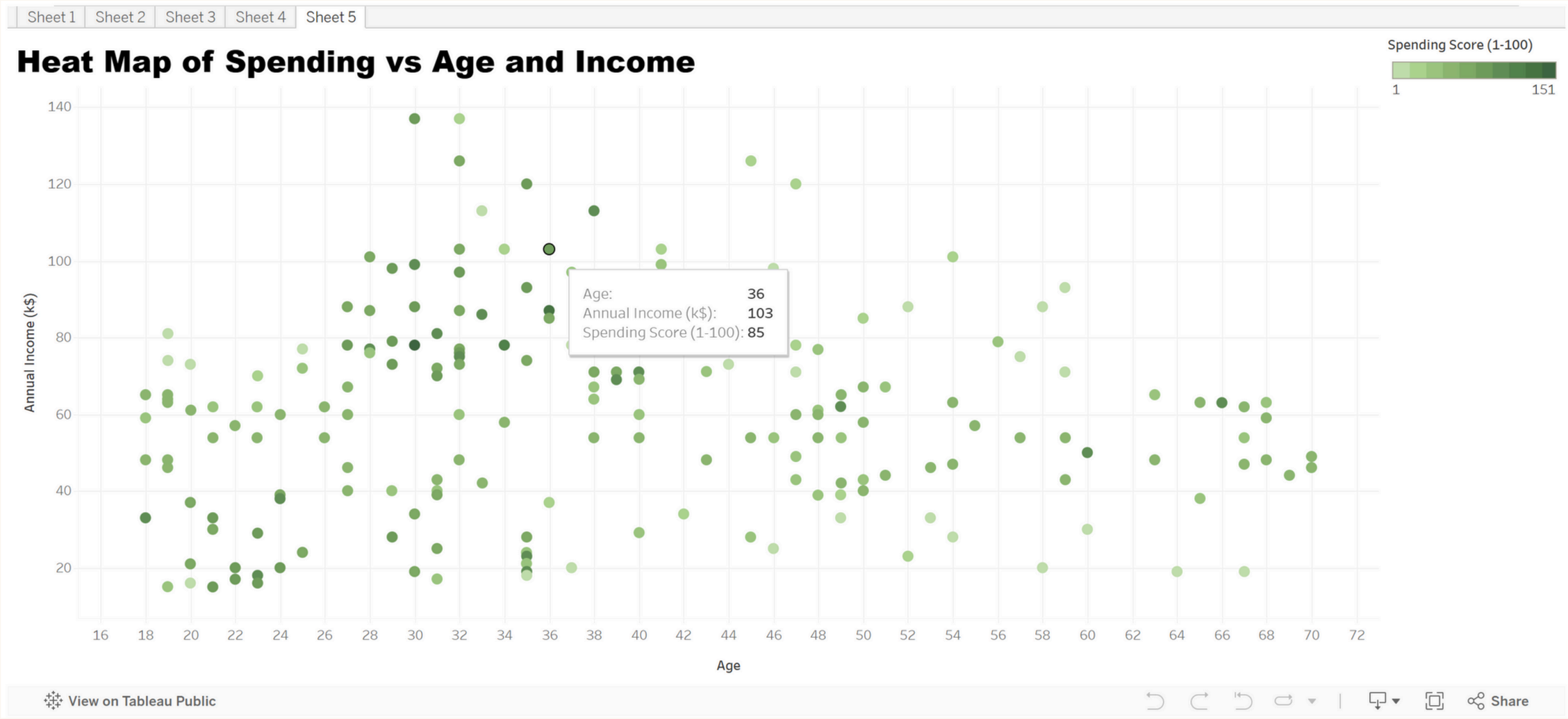
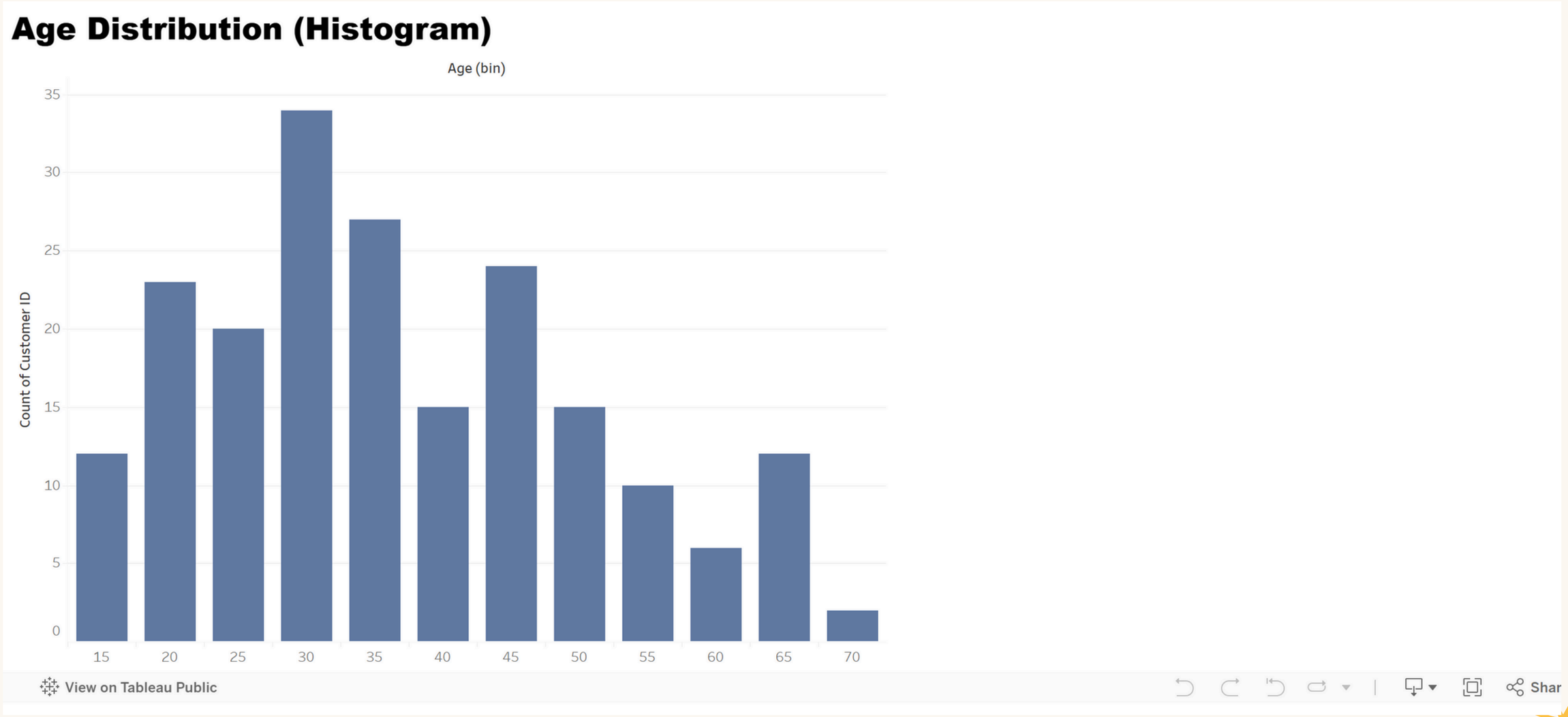
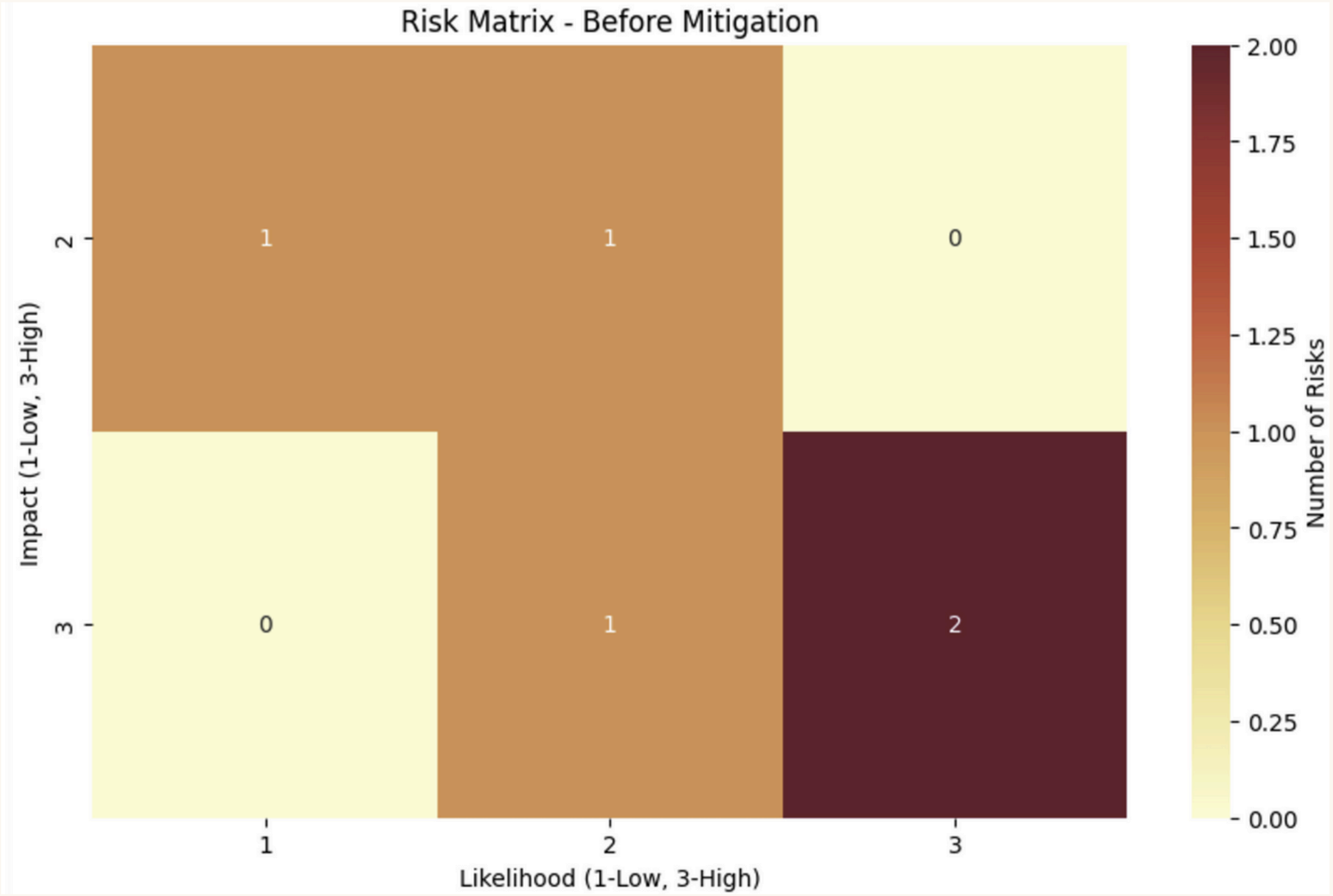
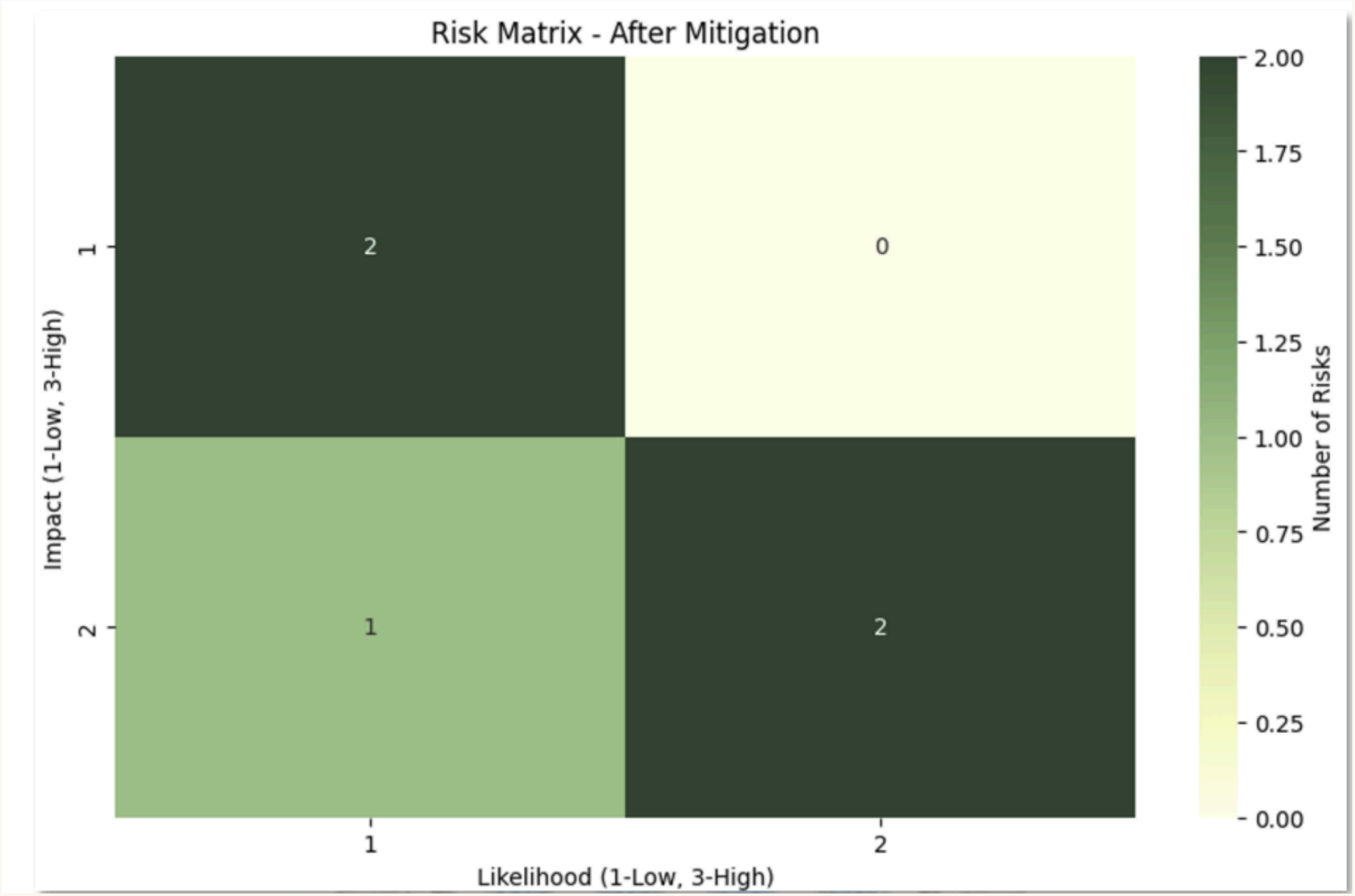


TABLEAU VISUALIZATION



VISUALIZING RISKS AND MITIGATION USING A RISK MATRIX




Risk Matrix - After Mitigation
Impact (1-Low, 3-High)
Likelihood (1-Low, 3-High)
Number of Risks
2
0
1
2
1
2
1
2
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

OVERALL ANALYSIS AND INSIGHTS(RISK ANALYSIS)

MITIGATION EFFECTIVENESS:

- THE VISUALIZATIONS CLEARLY SHOW THAT THE RISKS SCALED THE MITIGATION STRATEGIES YOU APPLIED WORKED AS RISKS SHIFTED FROM THE HIGH-PROBABILITY HIGH-IMPACT ZONES.
- THE IDEA OF MINIMIZING THE CHANCES AND VISIBILITY OF RISKY ITEMS IS QUITE ESSENTIAL TO SUSTAINING THE HEALTH OF THE PROGRAM AND IS SUGGESTIVE OF GOOD RISK MANAGEMENT AT THIS PLACE.

RISK CONTROL:

- CERTAIN HIGH-RISK ITEMS WERE EFFECTIVELY IDENTIFIED AND RELOCATED FROM THE AREAS OF HIGHEST CONCERN.
- HENCE THE BEFORE AND AFTER COMPARISONS ARE USEFUL IN ENSURING STAKEHOLDERS ARE CONVINCED THAT THE APPLIED MITIGATIONS ARE AS PLANNED.

REMAINING RISKS:

- THE TRANSITION WITHIN THE LIKELIHOOD DOMAIN OF HIGH-LIKELIHOOD ITEMS TO LOWER LIKELIHOOD CATEGORIES REVEALS PREVENTIVE ACTIONS THAT ARE DECREASING RISK EXPOSURE.

SUMMARY STATISTICS

PROJECT TIMELINE: 10 MONTHS

BUDGET: $1,200,000

PHASES OF THE PROJECT:

1. REQUIREMENTS GATHERING
2. DESIGN
3. DEVELOPMENT
4. TESTING
5. DEPLOYMENT

SUMMARY STATISTICS

CURRENT STATUS(LET'S ASSUME):

- MONTHS COMPLETED: 4 MONTHS OUT OF 10
- PLANNED VALUE (PV) AT THE END OF MONTH 4: $400,000
- EARNED VALUE (EV) AT THE END OF MONTH 4: $320,000
- ACTUAL COST (AC) AT THE END OF MONTH 4: $450,000

SUMMARY STATISTICS

KEY METRICS CALCULATION

1. COST VARIANCE (CV)

CV=EV-AC=320,000-450,000=-130,000

THE COST VARIANCE IS NEGATIVE, INDICATING COST OVERRUNS.

2.SCHEDULE VARIANCE (SV)

SV=EV-PV=320,000-400,000=-80,000

THE SCHEDULE VARIANCE IS ALSO NEGATIVE, INDICATING THAT THE PROJECT IS BEHIND SCHEDULE.

3.COST PERFORMANCE INDEX (CPI)

CPI=EVAC=320,000450,000≈0.71

A CPI OF LESS THAN 1 MEANS THE PROJECT IS OVER BUDGET.

4.SCHEDULE PERFORMANCE INDEX (SPI)

SPI=EVPV=320,000400,000=0.8

AN SPI OF LESS THAN 1 MEANS THE PROJECT IS BEHIND SCHEDULE.

5.ESTIMATE AT COMPLETION (EAC)

EAC=BUDGETCPI=1,200,0000.71≈1,690,141

THE PROJECTED COST OF COMPLETING THE PROJECT IS APPROXIMATELY $1,690,141, WHICH IS HIGHER THAN THE ORIGINAL BUDGET.

SUMMARY STATISTICS

ANALYSIS OF CURRENT PERFORMANCE

- **COST STATUS:** THE PROJECT IS OVER BUDGET, WITH COST OVERRUNS OF $130,000.
- **SCHEDULE STATUS:** THE PROJECT IS BEHIND SCHEDULE BY $80,000 IN TERMS OF THE VALUE OF WORK THAT SHOULD HAVE BEEN COMPLETED.
- **FUTURE PERFORMANCE:** WITH A CPI OF 0.71, THE PROJECT IS LIKELY TO CONTINUE EXPERIENCING COST OVERRUNS UNLESS CORRECTIVE ACTIONS ARE TAKEN. SIMILARLY, THE SPI OF 0.8 SUGGESTS A SIGNIFICANT DELAY IN THE PROJECT TIMELINE.
- **PROJECT STATUS:**
  - **COST:** OVER BUDGET
  - **SCHEDULE:** BEHIND SCHEDULE

REQUIREMENT ANALYSIS

MIRO BOARD LINK

From the User Perspective

FEASIBLE POINTS
- Enhances shopping experience
- Reduced need for store assistance
- Informed shopping decisions
- Access to promotions and deals

NON FEASIBLE POINTS
- Difficult accessing real-time data
- Privacy and data security concerns
- Lack of customization for different stores
- Lack of offline functionality

From the Employees Perspective

FEASIBLE POINTS
- Increased customer satisfaction
- Boost in sales
- Competitive advantage

NON FEASIBLE POINTS
-Potential impact on customer representative jobs due to reduced need for assistance

From the Mall Management & Retailers Perspective

FEASIBLE POINTS
- Ensuring store placements and product details are always current
- Product listings: Collaboration with developers to update inventory and product placements regularly

NON-FEASIBLE POINTS
- Retailers may need to provide regular information on deals and sales to attract customers (may depend on retailer cooperation)

CONCLUSIONS ON BASIS OF SUMMARY STATISTICS

CURRENT STATUS: THE PROJECT IS OVER BUDGET AND BEHIND SCHEDULE.

IMPACT ON DELIVERY: IF THE CURRENT TREND CONTINUES, THE PROJECT MAY NOT BE DELIVERED ON TIME OR WITHIN THE ALLOCATED BUDGET.

FINAL RECOMMENDATION: CONSIDER REALLOCATION OF RESOURCES, STRICTER BUDGET MANAGEMENT, AND POSSIBLY ADJUSTING THE SCOPE OR TIMELINE TO GET THE PROJECT BACK ON TRACK.

THANK YOU!

www.reallygreatsite.com