

GOWelfare

Provides an app for local purchases that allows employees to spend welfare credits at their favorite local shops and not just at large retail chains.

Aim of the Project:

Implement systems to optimize the welfare platform and maximise:

- profit
- user's satisfaction.

Task 1 – Load the Datasets

giftcard_model

Lists all the giftcards available on the welfare application, along with their associated brands.

giftcard_model.nunique()

id	96
amount	12
new_brand_name	33
brand_id	34
category	9
fee	17
num_punti_vendita	32

transaction

Lists all basic operations of the users towards the welfare platform.

transaction.nunique()

id	18409
amount	537
type	1
status	1
timestamp	18409
transaction_uuid_hash	18409
email_hash	15735
receiver_id_hash	15735

giftcard_obf

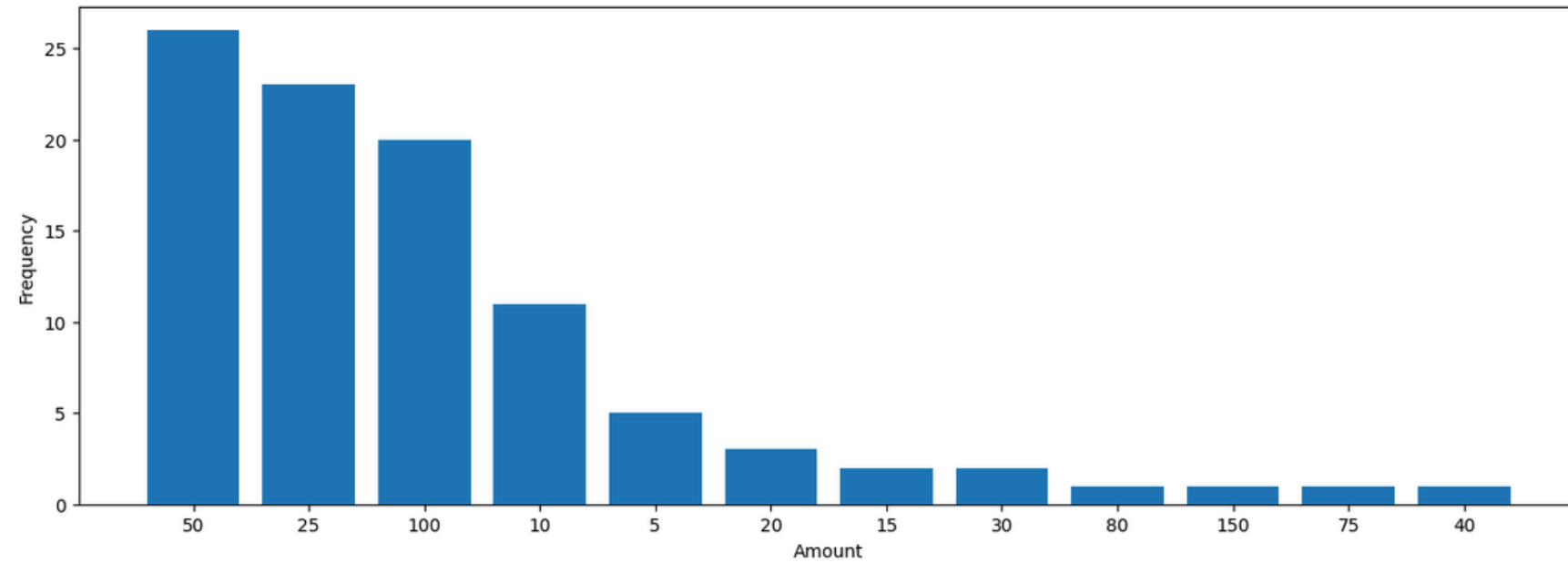
Lists all the giftcards bought by the users.

giftcard_obf.nunique()

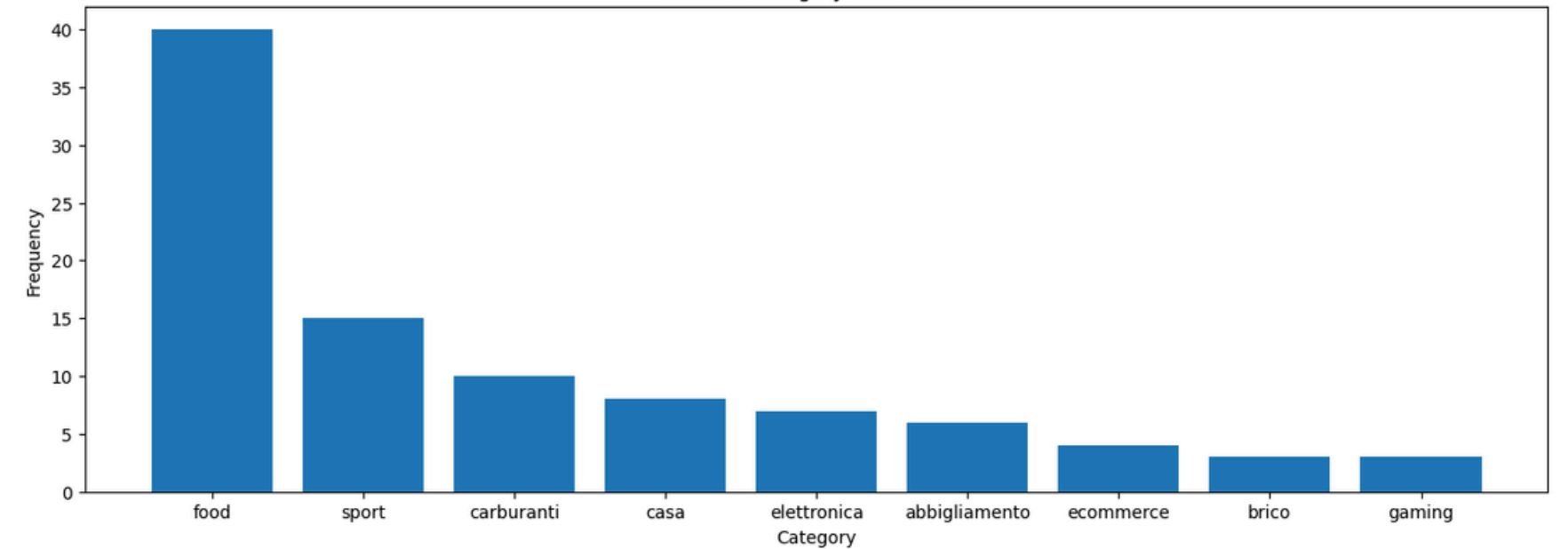
id	40189
provider_error_code	8
purchase_date	40189
giftcard_model_id	94
transaction_uuid_hash	40189
client_email_hash	9053

EDA on the Datasets

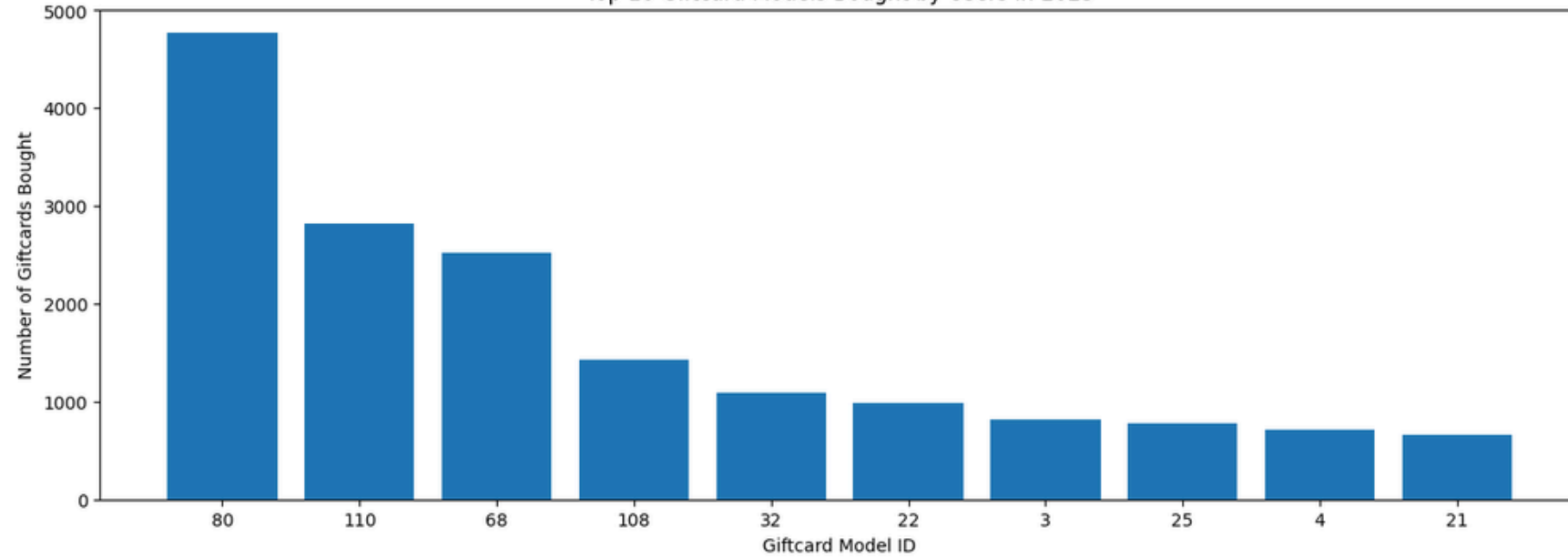
Giftcard Amount Distribution



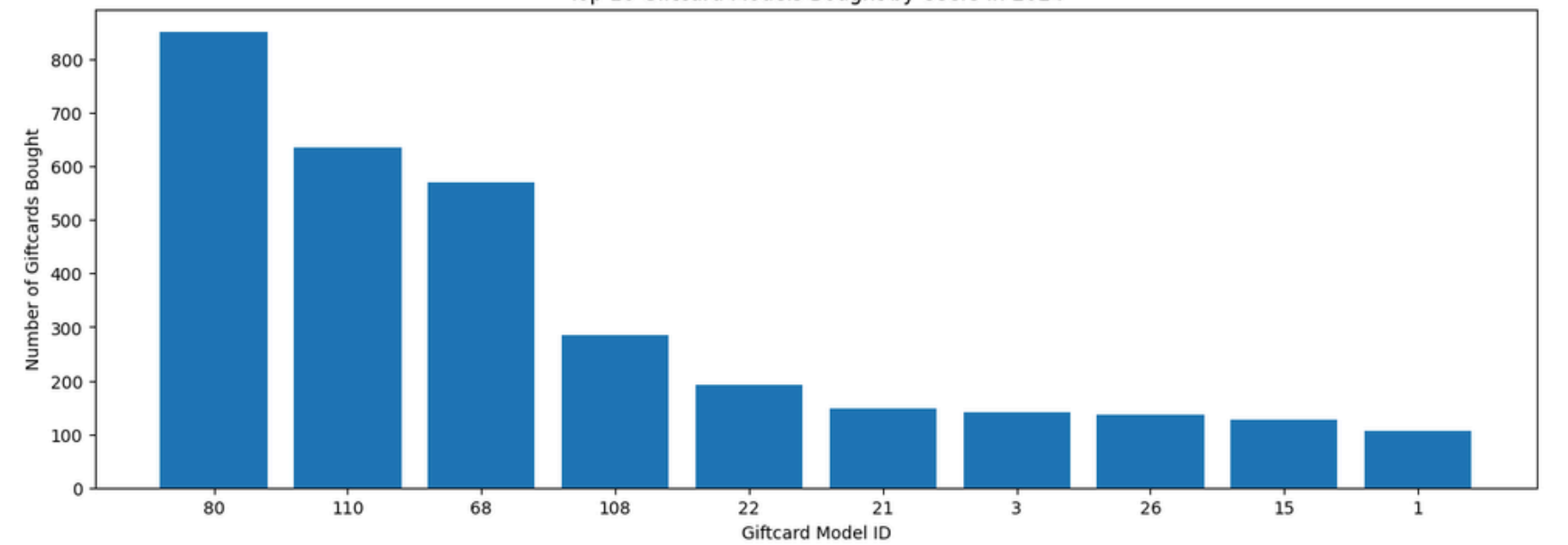
Giftcard Category Distribution



Top 10 Giftcard Models Bought by Users in 2023



Top 10 Giftcard Models Bought by Users in 2024



The board of directors, at the end of December 2023, want to know, for the FY 2023:

01 **Total Turnover:** € 2.299.656

02 **Total Spent:** € 1.002.055

03 **Total Profit:** € 26.645

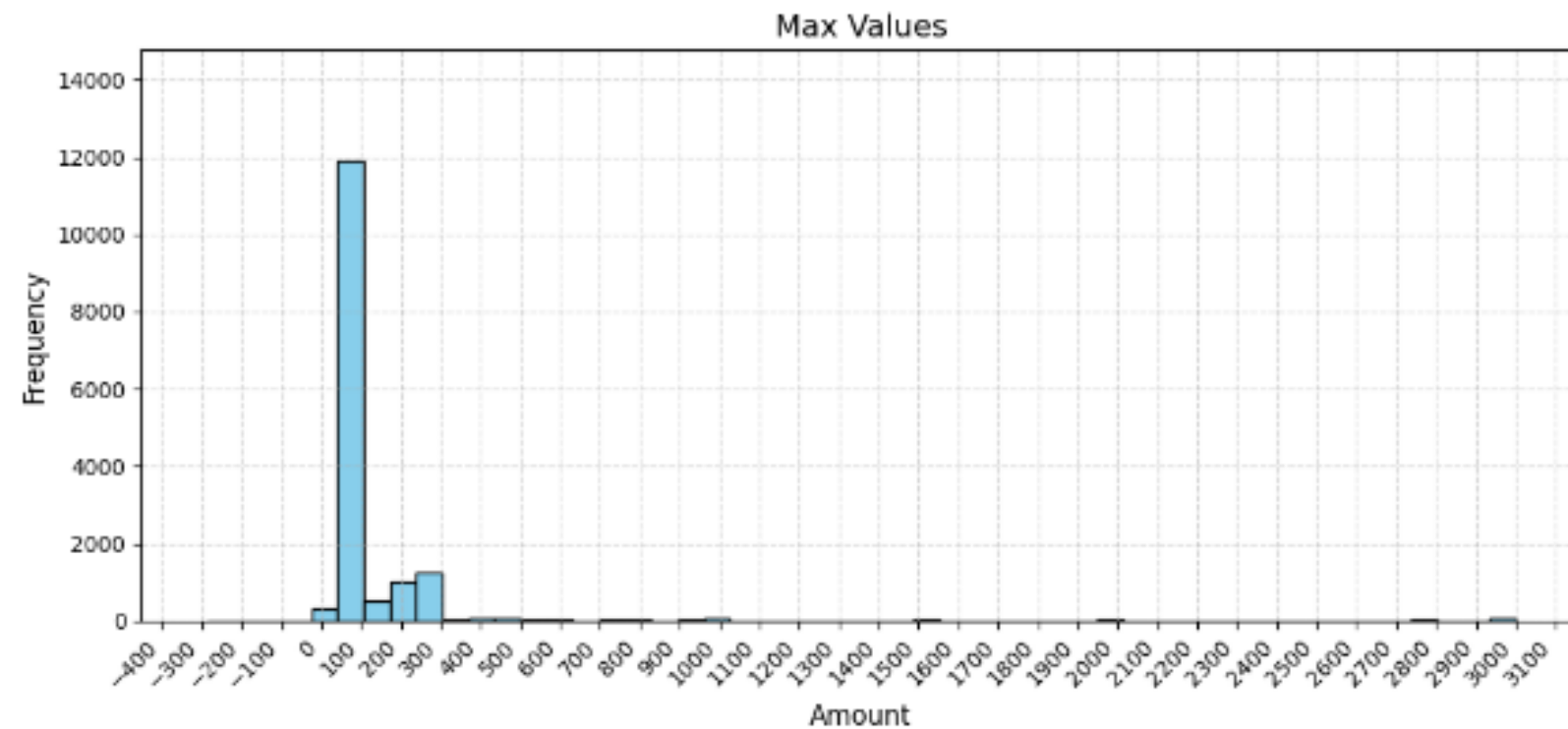
04 **Total Unspent:** € 1.297.601

05 **Profits if all the gift cards had as fee the maximum among the fees:** € 80.164

06 **Best performing brand and amount:**

- The brand that has **sold the most giftcards** is **CondividiAcquista**.
- The total amount gained from these giftcards is 95.420€ and the total number of the giftcards sold is 4771.
- The brand that has **earned the highest amount** is **SupermercatoElite**.
- The total amount earned from this brand is 125415€ and the total number of giftcards sold are 2424.

Distribution of User Credits

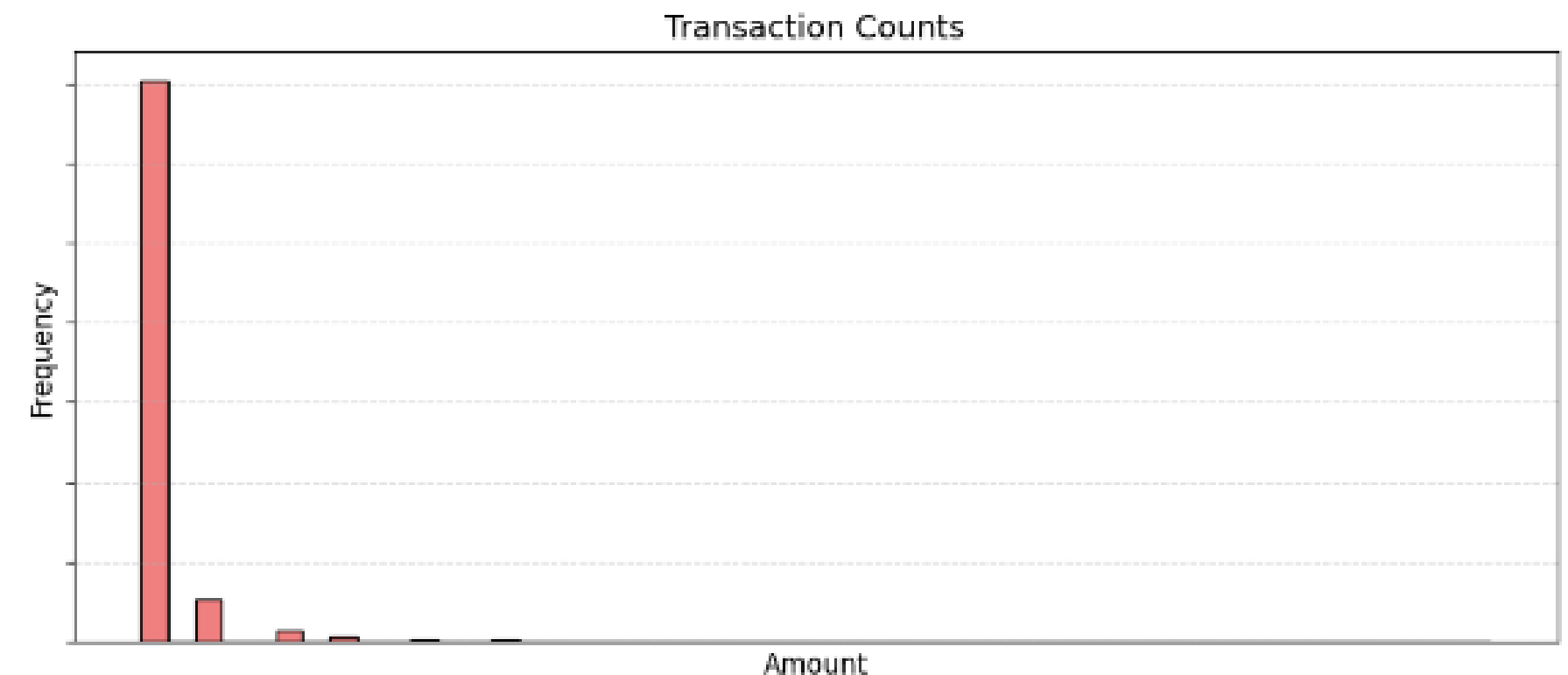


Histogram of Max Values:

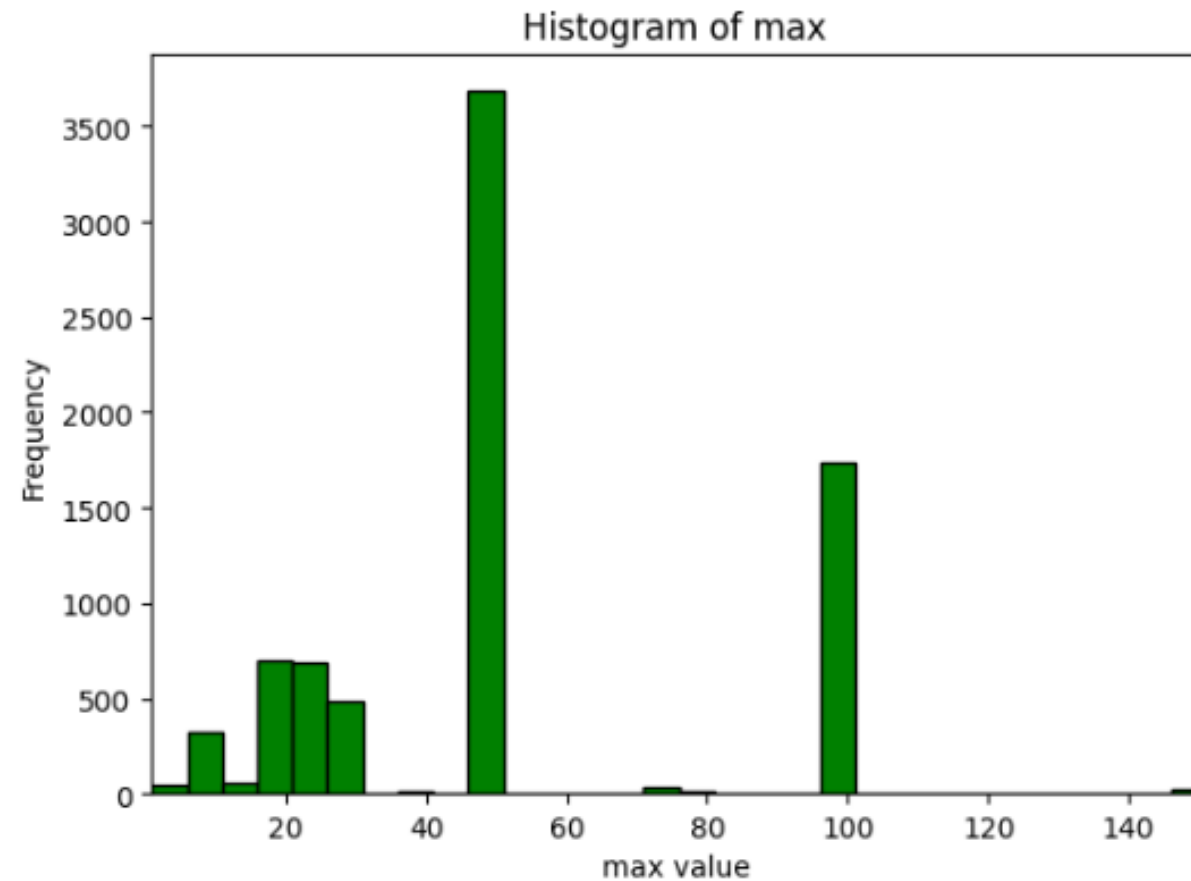
Indicates that the majority of users have maximum transaction values concentrated in a lower range, with few reaching higher values, suggesting infrequent high-value transactions.

Histogram of Transaction Counts:

Shows that most users conduct a low number of transactions, while a minority have a high count, suggesting a pattern of occasional engagement for most users and high activity for a few.



Distribution of Gift Cards Amounts

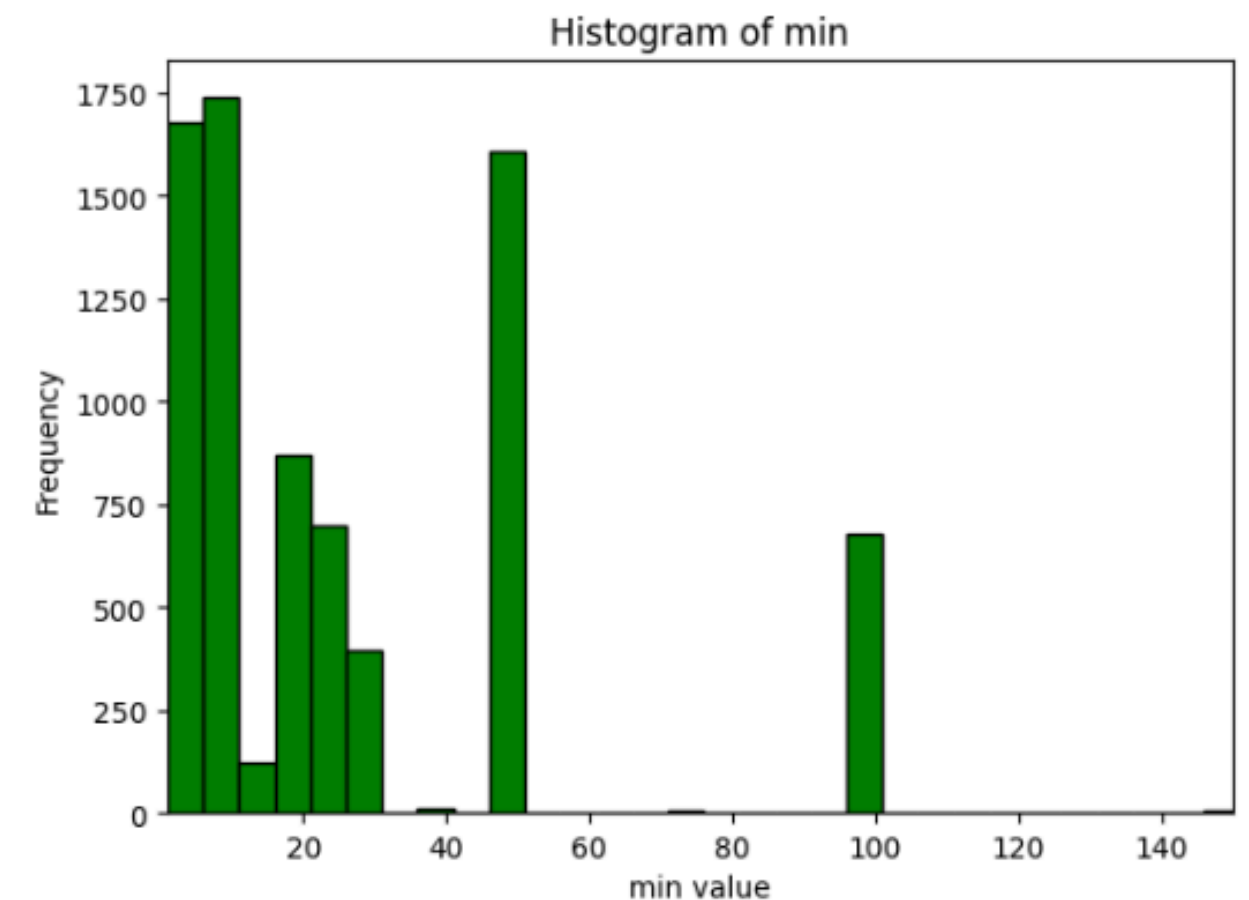


Histogram of Max Values:

This histogram shows that most users' highest purchase amount of gift cards is between 40 and 60, with over 4000 users buying gift cards in this range.

Histogram of Min Values:

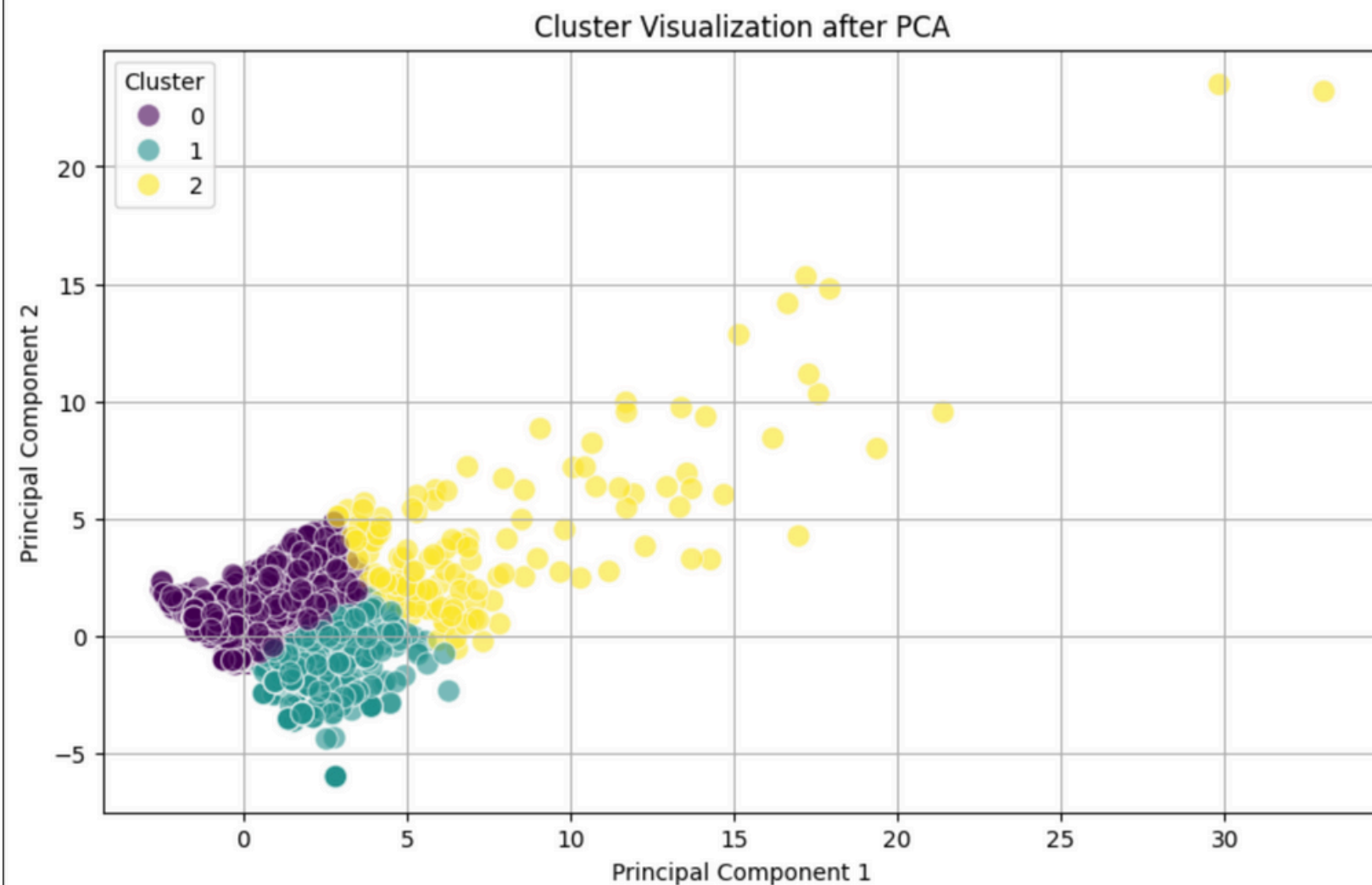
The histogram for minimum values is heavily skewed towards zero or low amounts, indicating that many users have purchased only small-value gift cards.



Users Similarity

Category Preferences:

- Cluster 2: Higher mean purchase count in food, e-commerce, electronics.
- Cluster 0: Lower mean purchase counts across all categories.
- Cluster 1: Moderate mean purchase count, slightly higher in e-commerce, food.



Transaction Statistics:

- Cluster 2: Higher mean transaction value, potentially higher spending behavior.

Observations:

- Cluster 2: Represents heavy spenders with higher mean and maximum transaction values.
- Cluster 0: Represents users with lower engagement.
- Cluster 1: Falls between the other two clusters, indicating moderate engagement.

Anomaly Detection

We identified outliers among users by calculating the interquartile range for each column and defining bounds for normal values, then flagging transactions falling outside these bounds as potential outliers. Then we compared the detected outliers's characteristics with the majority cluster's ones.

1. Size Disparity:

The majority cluster comprises 6148 users, while outliers are only 572, indicating a smaller but noteworthy subset with deviant behavior.

2. Transaction Amounts:

Majority cluster's mean transaction amount is moderate, contrasting with outliers' significantly higher mean, suggesting potential high-value transactions or irregular spending.

3. Transaction Range:

Majority cluster shows a consistent transaction range, whereas outliers exhibit a broader range, indicating diverse spending behaviors, from minimal to substantial spending.

4. Summary Statistics Comparison:

Outliers demonstrate disparities in summary statistics, including higher maximum transaction amounts, lower minimums (some negative), higher means, and wider transaction count ranges, highlighting distinct spending behaviors or anomalies.