

A woman with long dark hair, wearing a dark patterned top, is looking down at a tablet computer she is holding. The background is a blurred office or cityscape at night, with bokeh light effects. The entire image is overlaid with a semi-transparent blue filter. A solid dark blue vertical bar is on the far left.

CUSTOMER CHURN ANALYSIS

CONTENT SLIDE

- Introduction to Customer Churn.
- Data Collection and Preparation
- SQL for Data Extraction
- Python for Data Analysis
- Tableau for Data Visualization

The background of the slide is a dark blue overlay on a photograph. The photograph shows several hands of different people pointing at a large, light-colored map or blueprint spread out on a table. The hands are in various positions, some pointing directly at the camera, others pointing towards different parts of the map. The map itself has some faint lines and text, but it's not legible. The overall scene suggests a collaborative meeting or a presentation of a plan.

INTRODUCTION TO CUSTOMER CHURN

UNDERSTANDING CUSTOMER CHURN



Definition of Churn

Churn refers to the loss of customers or clients within a specific timeframe, often measured as a percentage of total customers at the start of that period.



Importance of Churn Analysis

Analyzing churn is crucial for businesses as it helps identify factors leading to customer loss and informs strategies to improve customer retention and engagement.



Industry Impact

Different industries experience varying rates of churn, which influences strategic decisions regarding customer acquisition, service improvements, and overall market competitiveness.

KEY METRICS IN CHURN ANALYSIS



Churn Rate Calculation

Churn rate is calculated by dividing the number of customers lost during a certain period by the number of customers at the beginning of that period, multiplied by 100.



Customer Lifetime Value (CLV)

CLV is a predictive metric that estimates the total revenue a business can expect from a single customer account over the entire duration of their relationship.



Retention Rate

Retention rate measures the percentage of customers who remain with a company over a specified timeframe, highlighting the effectiveness of customer engagement initiatives.

The background of the slide is a dark blue overlay on a photograph. The photograph shows several people's hands and forearms gathered around a large table, pointing at a complex data table or map. The data table has multiple columns and rows, with some cells containing numbers and others containing text. The overall scene suggests a collaborative data analysis or planning session.

DATA COLLECTION AND PREPARATION

DATA SOURCES



Internal Databases

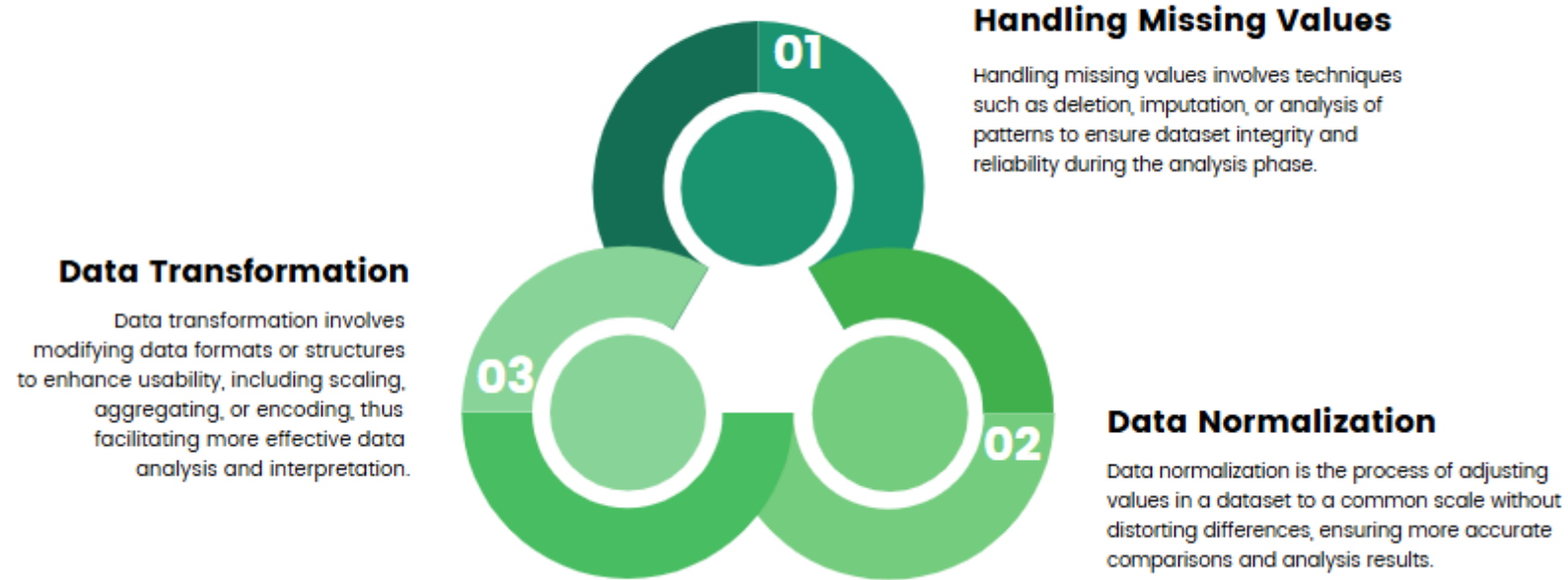
Internal databases are repositories within an organization that store data gathered from its operations, providing valuable insights for analysis and decision-making.



External Data Providers

External data providers offer datasets from outside the organization, including market research firms or public data sets, enhancing the internal data with broader context and trends.

DATA CLEANING



The background of the slide features a photograph of several hands pointing at a large document or map spread out on a table. The image is dimmed and has a solid blue color overlay applied to it. The text 'SQL FOR DATA EXTRACTION' is written in a white, serif font, positioned on the left side of the image.

SQL FOR DATA EXTRACTION

SQL BASICS



■ Select Statements

Select statements are fundamental SQL commands used to query data from a database, allowing users to specify columns, filter records, and order results effectively.

■ Aggregations and Joins

Aggregations in SQL enable the summarization of data using functions like COUNT, SUM, AVG, while joins combine records from two or more tables based on related columns.

SELECT, FROM, WHERE, JOIN, GROUP BY

-- Reterive all customers from north america region

```
select FirstName, LastName, customerid, region
from customers
where region='North America';
```

-- Number of customers who have taken annual subscrptions

```
select * from subscriptions;
select * from customers;
select count(*) as Total_subscriptions
from subscriptions
where PlanType='Annual';
```

- Retrieve the major reason fot the customer churn

```
select Reason, count(customerID) from churn
group by reason;
```

- Retrieve the customer details along with subscription details

```
select
c.CustomerID, c.FirstName, c.LastName, s.Subscript
ionID, s.PlanType
from customers c
Join subscriptions s on c.CustomerID =
s.CustomerID;
```



PYTHON FOR DATA ANALYSIS

PYTHON LIBRARIES



Pandas

Pandas is an essential library for data manipulation and analysis, providing data structures like DataFrames for handling structured data efficiently.



NumPy

NumPy is the foundational package for numerical computing in Python, offering support for multi-dimensional arrays and a variety of mathematical functions.



SciPy

SciPy builds on NumPy and provides additional functionality for scientific computing, including modules for optimization, integration, and statistical analysis.

DATA VISUALIZATION IN PYTHON

01



Matplotlib

Matplotlib is a versatile plotting library that enables users to create static, animated, and interactive visualizations in Python, making data interpretation easier.



02



Seaborn

Seaborn is a statistical data visualization library based on Matplotlib, offering a higher-level interface for drawing attractive and informative graphics.



DATA VISUALIZATION IN PYTHON

[31]: #5. Visualization of Active & Inactive Customers

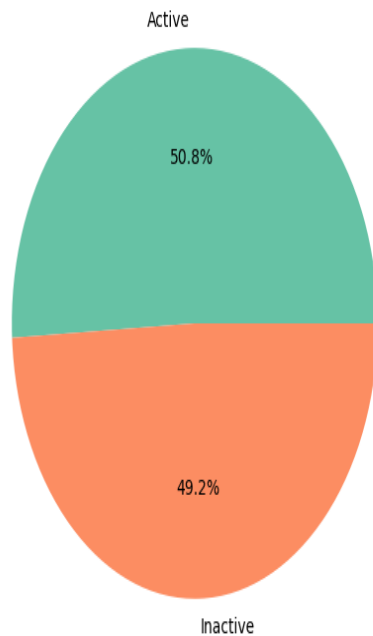
```
#Calculate the number of active / Inactive customers
```

```
status_count=customers_df['Status'].value_counts()
```

```
#Create a pie chart to show the distribution
```

```
plt.figure(figsize=(8,6))
plt.pie(status_count, labels=status_count.index, autopct='%1.1f%%', colors=['#66c2a5','#fc8d62'])
plt.title('Distribution of Customer Status')
plt.show()
```

Distribution of Customer Status



[39]: #7. Create a visualisation(bar chart) who joined after and before 2021

```
#converting JoinDate into DATE TIME FORMAT
```

```
customers_df['JoinDate']=pd.to_datetime(customers_df['JoinDate'],format='%d-%m-%Y')
```

```
# Filtering the customers who joined before and after 2021
```

```
before_2021=customers_df[customers_df['JoinDate']<'2021-01-01']
```

```
after_2021=customers_df[customers_df['JoinDate']>'2021-01-01']
```

```
# Count the number of customers in each of the category
```

```
counts={'Before 2021': len(before_2021), 'After 2021': len(after_2021)}
```

```
# create a bar chart
```

```
plt.bar(counts.keys(),counts.values(), color=['blue','green'])
```

```
plt.xlabel('Join Date Category')
```

```
plt.ylabel('No. of customers')
```

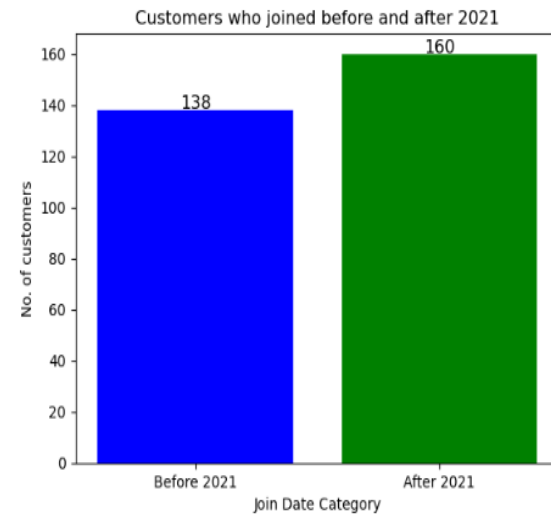
```
plt.title('Customers who joined before and after 2021')
```

```
#Add the exact numbers on the top of the bars
```

```
for i, (key,value) in enumerate(counts.items()):
    plt.text(i,value+0.5, str(value), ha='center', fontsize=12)
plt.show()
```

```
# Using for Loop, the Loop goes through each bar (Before 2021 and After 2021)
```

```
#For each bar, it places the customer count(value) as a Label slightly above the bar, position it at centre
#counts.items(): This gives you both the key and the values from dictionary
```



DATA VISUALIZATION IN PYTHON

```
[44]: #8. List down the active customers from Europe
Active_cx=customers_df[(customers_df['Region']=='Europe') & (customers_df['Status']=='Active')]
result=len(Active_cx)
print(result)
```

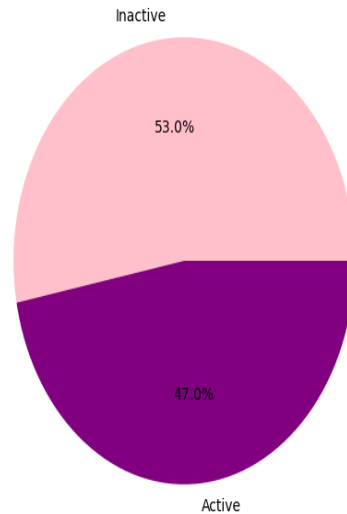
47

```
[49]: #Visualisation of active and Inactive Customers in Europe.
# Filtering the european customers
customers_europe=customers_df[customers_df['Region']=='Europe']

#Count the number of active and Inactive
status_counts= customers_europe['Status'].value_counts()

plt.figure(figsize=(8,6))
plt.pie(status_counts, labels=status_counts.index, autopct='%1.1f%%',colors=['pink','purple'])
plt.title('Distribution of Customers')
plt.show()
```

Distribution of Customers



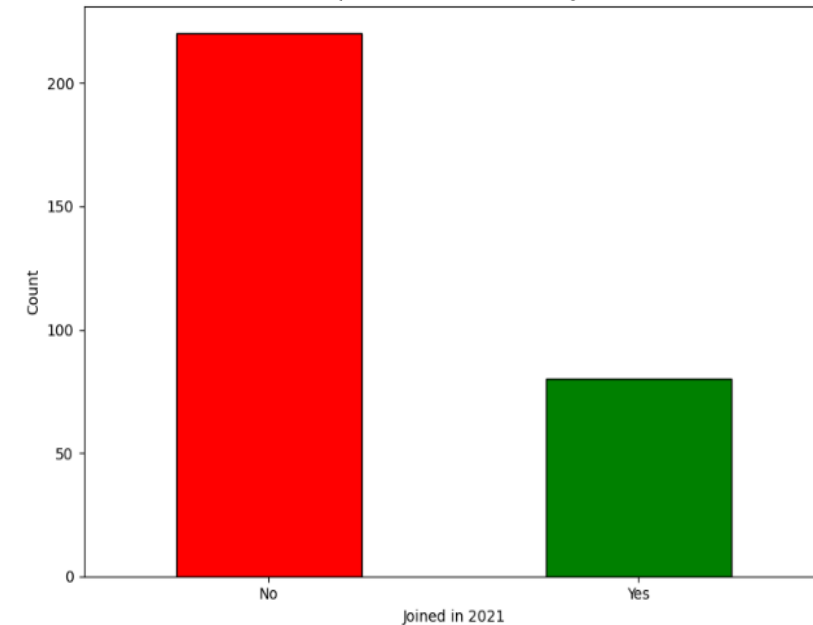
```
.01]: customers_df['JoinDate']=pd.to_datetime(customers_df['JoinDate'],format='%d-%m-%Y')
customers_df['Joined2021'] = pd.to_datetime(customers_df['JoinDate']).dt.year == 2021
```

```
# Count the number of customers who are in yes or no
joined_2021_counts = customers_df['Joined2021'].value_counts()
```

```
# Create a bar chart
plt.figure(figsize=(8, 6))
joined_2021_counts.plot(kind='bar', color=['red', 'green'], edgecolor='black')
plt.xlabel('Joined in 2021')
plt.ylabel('Count')
plt.title('Bar Chart Comparison of Customers Who Joined in 2021')
plt.xticks(ticks=[0, 1], labels=['No', 'Yes'], rotation=0)
plt.tight_layout()
```

```
plt.show()
```

Bar Chart Comparison of Customers Who Joined in 2021



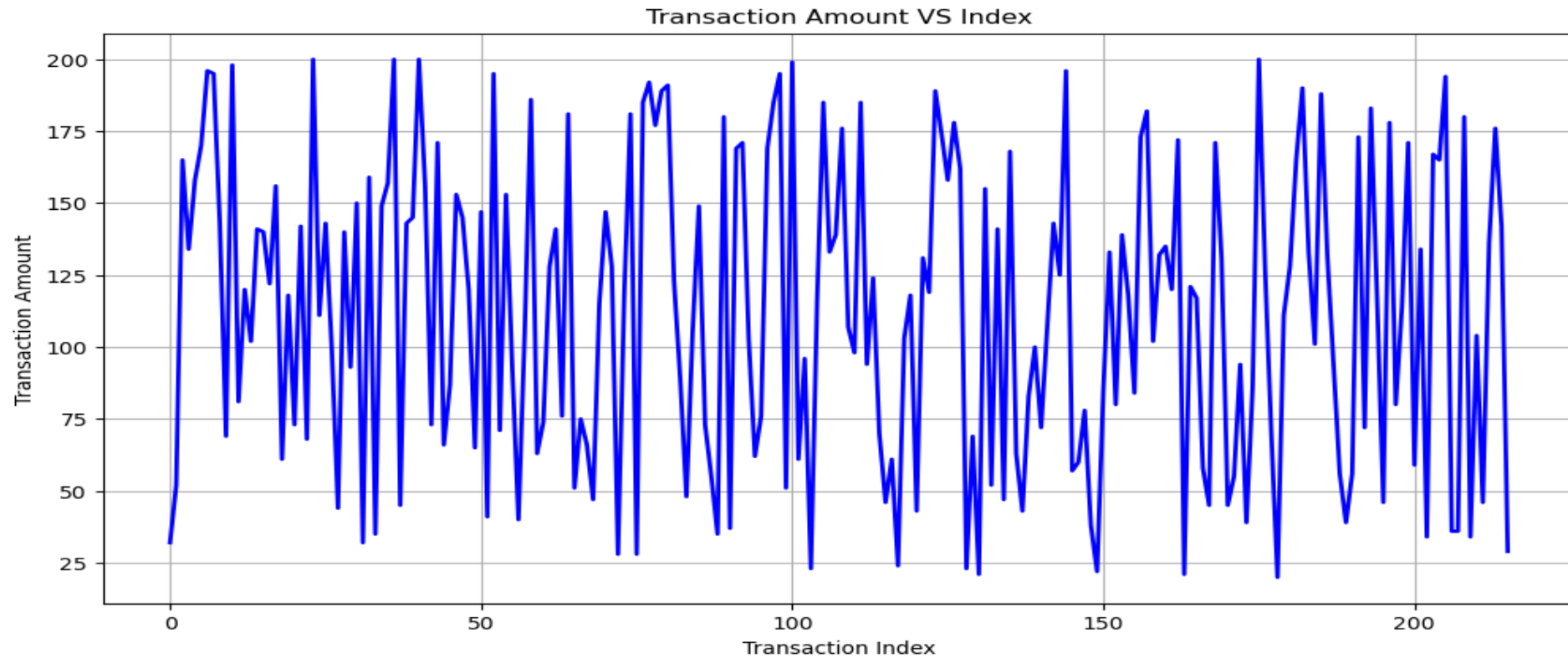
DATA VISUALIZATION IN PYTHON

```
#calculate the average transaction amount
Average_amount= Transactions_df['Amount'].mean()
print(f"Average Transaction Amount: {Average_amount}")
```

Average Transaction Amount: 112.19907407407408

```
# Visualize / Analysis for Transaction Amount w.r.t to Index - Line graph
plt.figure(figsize=(12, 6))
plt.plot(Transactions_df.index, Transactions_df['Amount'],linestyle='-', linewidth=2, color='blue')

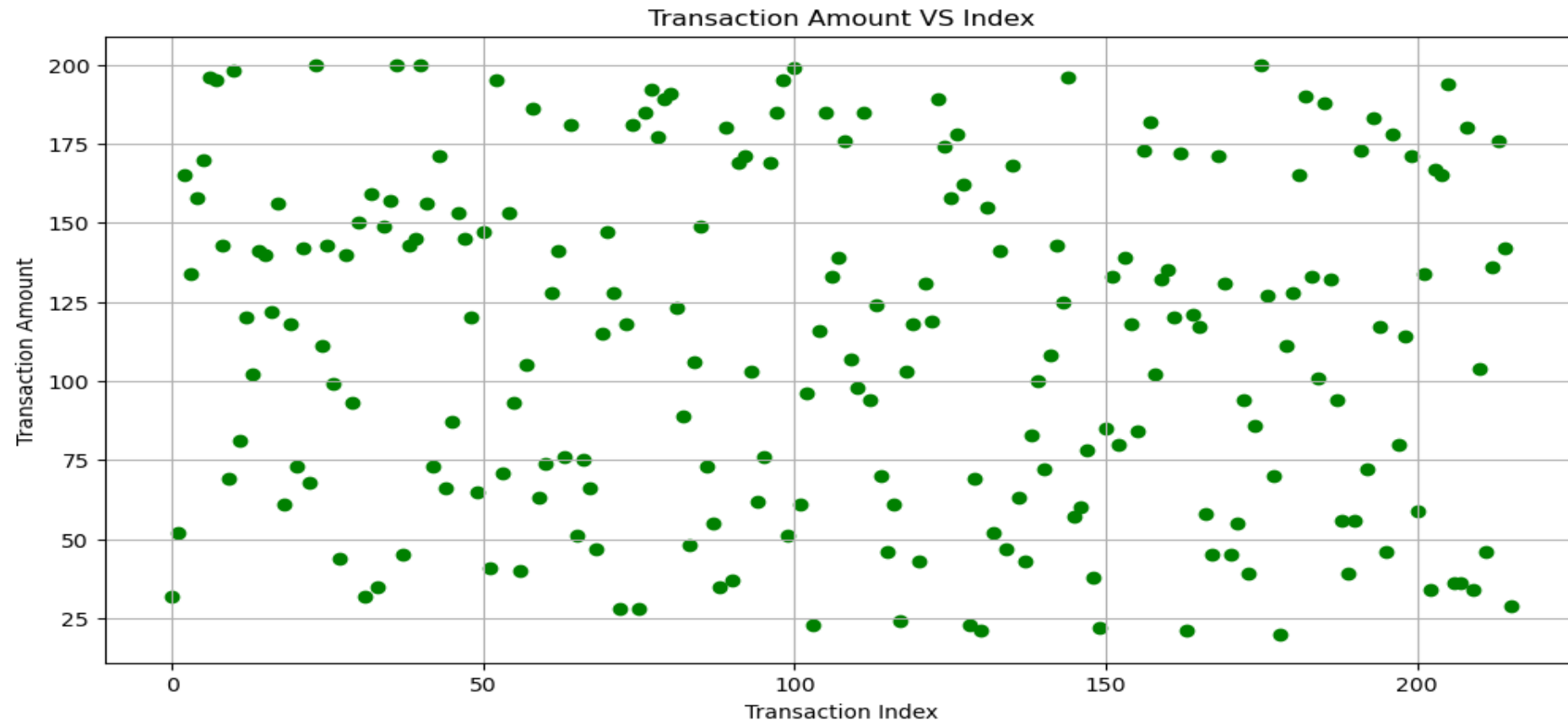
plt.title('Transaction Amount VS Index')
plt.xlabel('Transaction Index')
plt.ylabel('Transaction Amount')
plt.grid(True)
plt.show()
```



DATA VISUALIZATION IN PYTHON

```
# Visualize / Analysis for Transaction Amount w.r.t to Index - scatter plot
plt.figure(figsize=(12, 6))
plt.scatter(Transactions_df.index, Transactions_df['Amount'],color='Green')

plt.title('Transaction Amount VS Index')
plt.xlabel('Transaction Index')
plt.ylabel('Transaction Amount')
plt.grid(True)
plt.show()
```



The background of the slide is a dark blue overlay on a photograph. The photograph shows several people's hands and forearms gathered around a large table. They appear to be pointing at and discussing a large-scale data visualization or map spread across the table. The lighting is somewhat dim, and the focus is on the collaborative interaction with the data.

TABLEAU FOR DATA VISUALIZATION

INTRODUCTION TO TABLEAU

Connecting Data Sources

Tableau allows users to connect to various data sources like Excel, SQL databases, and cloud data, enabling seamless integration and analysis across multiple platforms.

Data Preparation in Tableau

Data preparation involves cleaning, transforming, and organizing data within Tableau to ensure accurate visualizations, facilitating effective analysis and decision-making processes.



CREATING DASHBOARDS



Interactive Elements

Interactive elements such as filters, drop- downs, and buttons enhance user engagement, allowing viewers to manipulate visualizations for personalized insights and deeper data exploration.



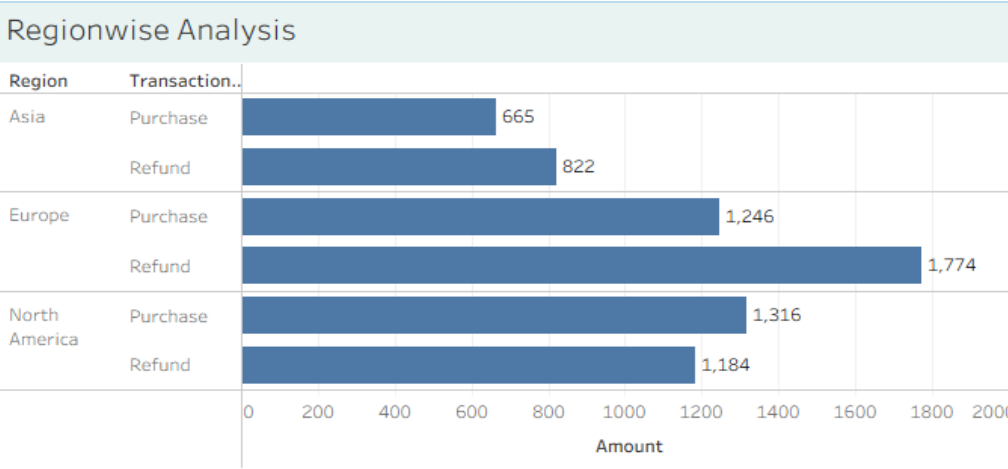
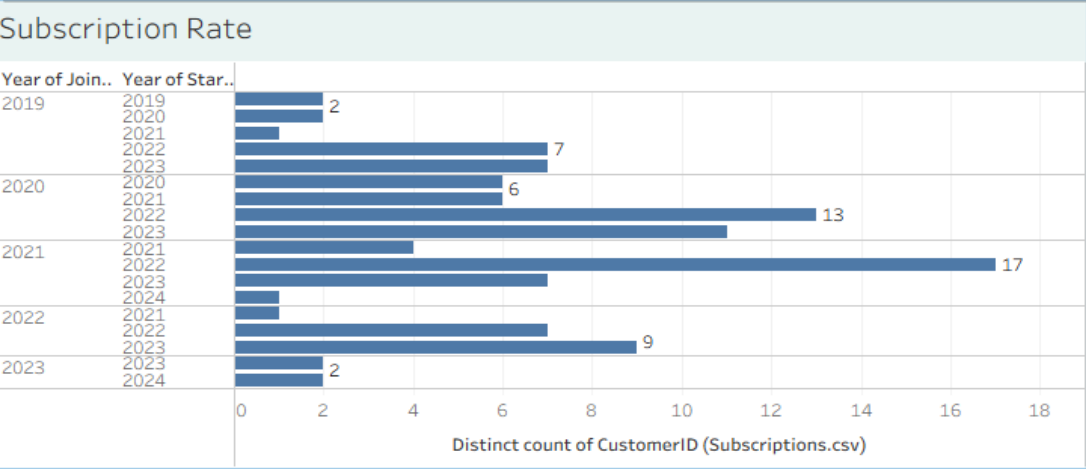
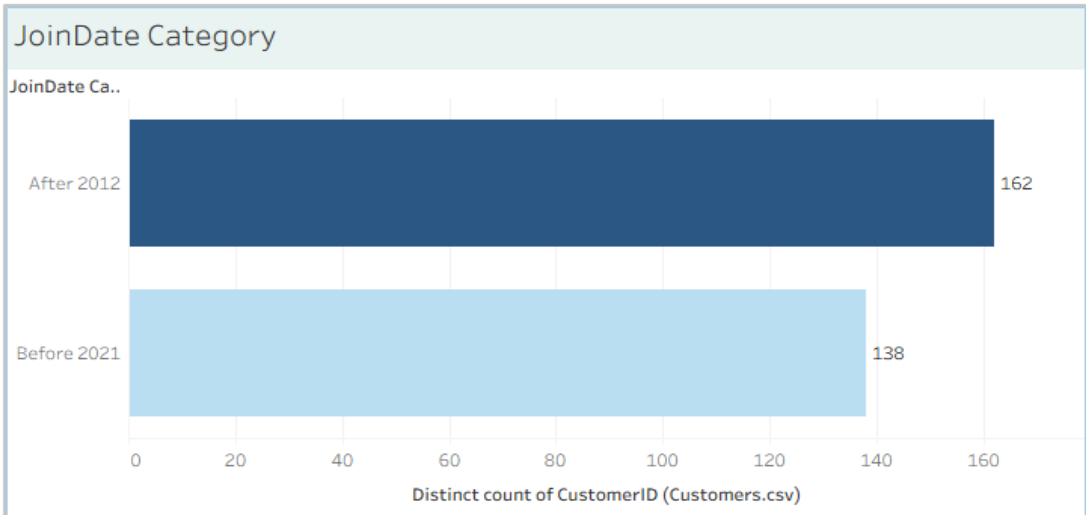
Real-time Updates

Real- time updates enable dashboards to present current data dynamically, ensuring that users have access to the latest information for timely decision- making and strategic initiatives.

VISUALIZING CHURN ANALYSIS REPORT

Customers Churn Analysis

Distinct count of CustomerID (Customers.csv)



THANKS

