Optimizing Customer Satisfaction: A Deep Dive into Spar Nord Bank ATM Transactions

- The Pie Sparkers
- Bhavana Prasad, Jay Dale, Rashmi Shree , Sawan Beli

Introduction

Introduction

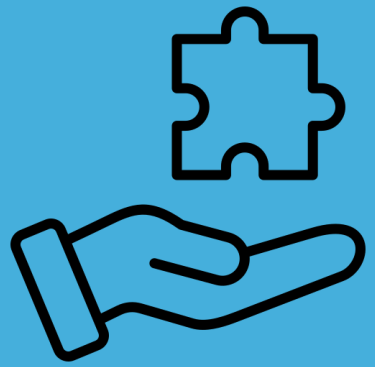
- Spar Nord Bank is a Danish financial institution that aims to improve its ATM management system and customer experience using data analysis.
- The project involves using different AWS services to collect and analyze transactional data from the bank's ATMs to identify customer behavior patterns and trends, make suggestions for improving the ATM network, and establish best practices for data governance and privacy.
- The project's ultimate goal is to help Spar Nord Bank optimize its services, reduce costs and risks, and provide a better overall experience for its customers by making data-driven decisions based on accurate, consistent, and secure data.

Goals and Objectives

Goals and Objectives

- To analyze Spar Nord Bank's ATM transaction data to gain insights into customer behavior and preferences.
- To identify patterns and trends in ATM usage, such as peak usage times and popular transaction types.
- To explore the relationship between weather conditions and ATM usage, and how this might inform the bank's approach to managing its ATM network.
- To develop effective data visualization techniques to communicate the findings of the analysis to key stakeholders within the bank.
- To make recommendations for future improvements to the ATM network, based on the insights gained from the analysis.

Problem and Motivation



Problem and Motivation

- The Spar Nord Bank has a large ATM network, but there is currently a lack of insights into how customers use the network. This makes it difficult for the bank to optimize its ATM services and improve the customer experience.
- With the increasing popularity of mobile banking and other digital services, it is important for the Spar Nord Bank to understand how ATM usage is evolving and how they can keep their ATM network relevant and useful to customers.
- By analyzing ATM transaction data, the Spar Nord Bank can gain valuable insights into customer behavior, preferences, and needs.
- This can help the bank make data-driven decisions about where to allocate resources and how to improve its ATM services.

Data Overview

Data Overview

- The data used in this project is transactional data of Spar Nord Bank's ATM network from Kaggle. The dataset contains over 2.5 million records of ATM transactions conducted by customers of Spar Nord Bank.

- The data includes information about the time and location of each transaction, as well as the type of transaction and the card used. Dataset samples were used to provide a glimpse of the actual data and its structure.

- These samples were used to get an idea of the data's format, the types of information included, and to help decide which data cleaning techniques to apply.

Sample Dataset

	year	month	day	weekday	hour	atm_status	atm_id	atm_manufacturer	atm_location	atm_streetname	...	temp	pressure	humidity	wind_speed	wind_
0	2017	January	1	Sunday	0	Active	1	NCR	Næstved	Farimagsvej	...	281.15	1014	87	7	
1	2017	January	1	Sunday	0	Inactive	2	NCR	Vejgaard	Hadsundvej	...	280.64	1020	93	9	
2	2017	January	1	Sunday	0	Inactive	2	NCR	Vejgaard	Hadsundvej	...	280.64	1020	93	9	
3	2017	January	1	Sunday	0	Inactive	3	NCR	Ikast	Rådhusstrædet	...	281.15	1011	100	6	
4	2017	January	1	Sunday	0	Active	4	NCR	Svogerslev	Brønsager	...	280.61	1014	87	7	
5	2017	January	1	Sunday	0	Active	5	NCR	Nibe	Torvet	...	280.64	1020	93	9	
6	2017	January	1	Sunday	0	Active	6	NCR	Fredericia	Sjællandsgade	...	281.15	1014	93	7	
7	2017	January	1	Sunday	0	Active	7	Diebold Nixdorf	Hjallerup	Hjallerup Centret	...	280.64	1020	93	9	
8	2017	January	1	Sunday	0	Active	8	NCR	Glyngøre	Færgevej	...	281.15	1011	100	6	
9	2017	January	1	Sunday	0	Active	9	Diebold Nixdorf	Hadsund	Storegade	...	280.64	1020	93	9	

10 rows × 33 columns

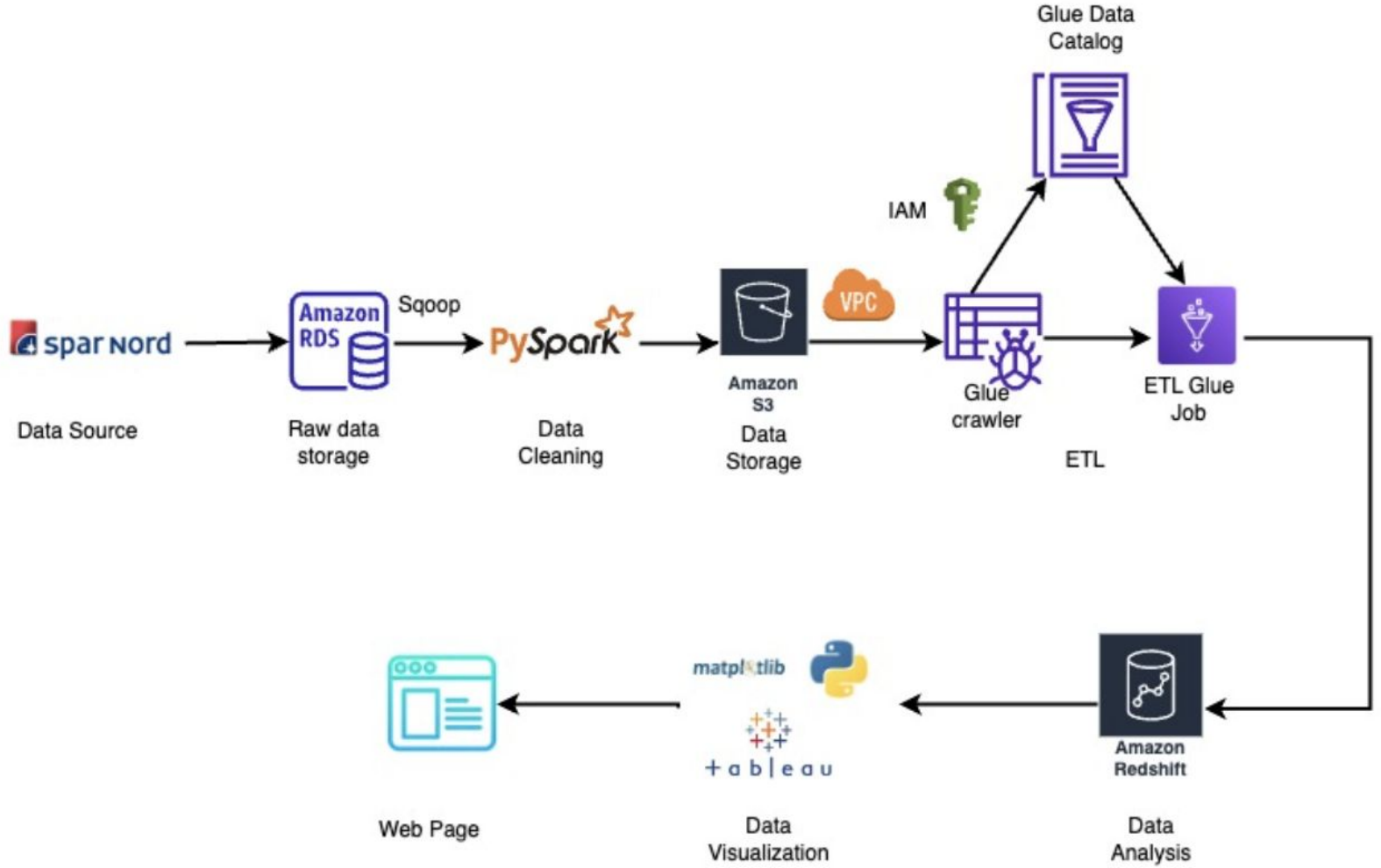
Architecture and Project Flow

Architecture and Project Flow

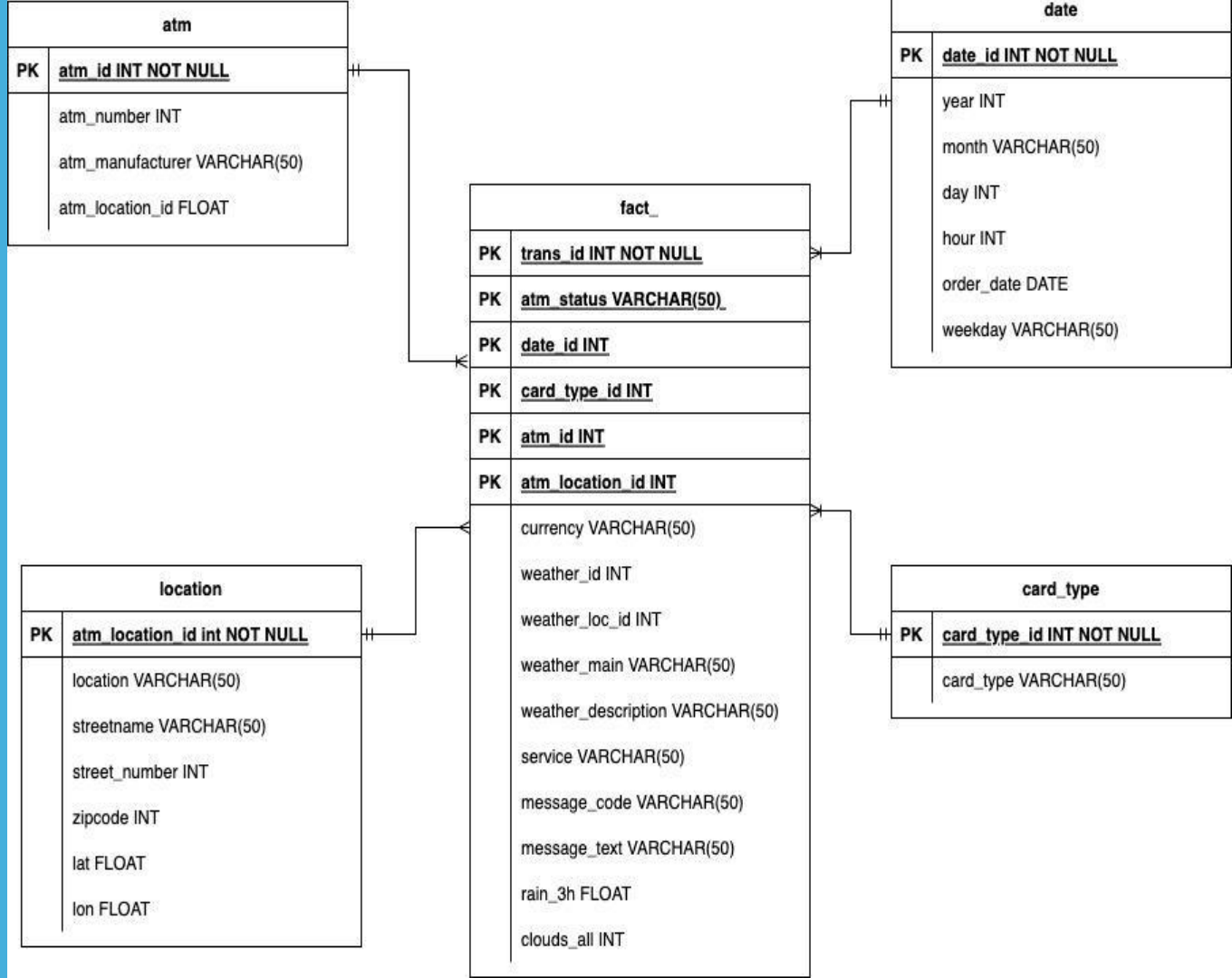
- The project architecture consisted of several components such as data engineering, data analytics, and data visualization. The data was initially imported from the RDS using Sqoop and then processed using an ETL pipeline.

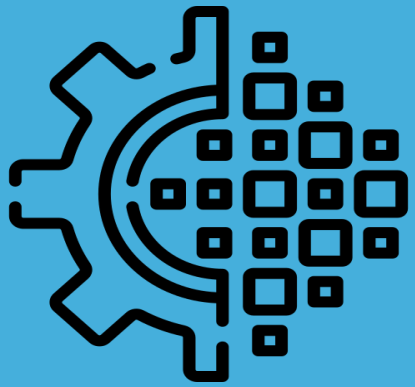
- The data modeling process involved designing an entity-relationship (ER) diagram to illustrate the relationships between the various data entities in the dataset. This helped to create a clear understanding of the data flow and how different elements were related to one another.

- The project flow involved a series of steps, including importing the data, cleaning and transforming the data, storing it in the data warehouse, analyzing the data, and finally, visualizing the data. Each step was carefully designed to ensure that the data was processed accurately and efficiently, and that the end results were reliable and insightful.


Glue Data Catalog
IAM
spar nord
Amazon RDS
Sqoop
PySpark
VPC
Amazon S3
Glue crawler
ETL Glue Job
Data Source
Raw data storage
Data Cleaning
Data Storage
ETL
matplotlib
tableau
Amazon Redshift
Web Page
Data Visualization
Data Analysis

Data Model


atm
PK atm_id INT NOT NULL
atm_number INT
atm_manufacturer VARCHAR(50)
atm_location_id FLOAT
date
PK date_id INT NOT NULL
year INT
month VARCHAR(50)
day INT
hour INT
order_date DATE
weekday VARCHAR(50)
fact_
PK trans_id INT NOT NULL
PK atm_status VARCHAR(50)
PK date_id INT
PK card_type_id INT
PK atm_id INT
PK atm_location_id INT
currency VARCHAR(50)
weather_id INT
weather_loc_id INT
weather_main VARCHAR(50)
weather_description VARCHAR(50)
service VARCHAR(50)
message_code VARCHAR(50)
message_text VARCHAR(50)
rain_3h FLOAT
clouds_all INT
location
PK atm_location_id int NOT NULL
location VARCHAR(50)
streetname VARCHAR(50)
street_number INT
zipcode INT
lat FLOAT
lon FLOAT
card_type
PK card_type_id INT NOT NULL
card_type VARCHAR(50)



Data Engineering

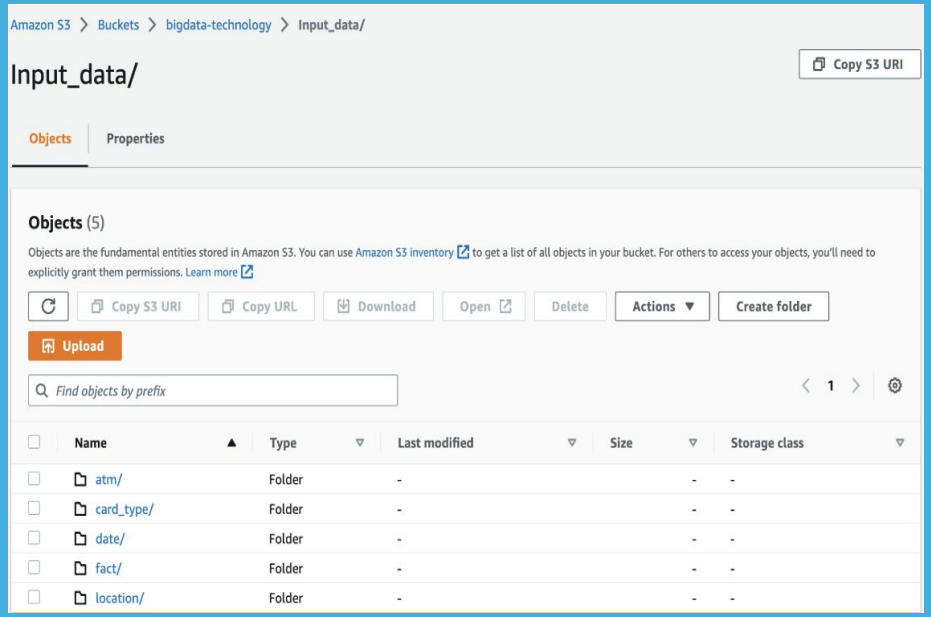
Data Engineering

- **Importing Data: The Spar Nord bank ATM transactions data is extracted using Sqoop from RDS.**

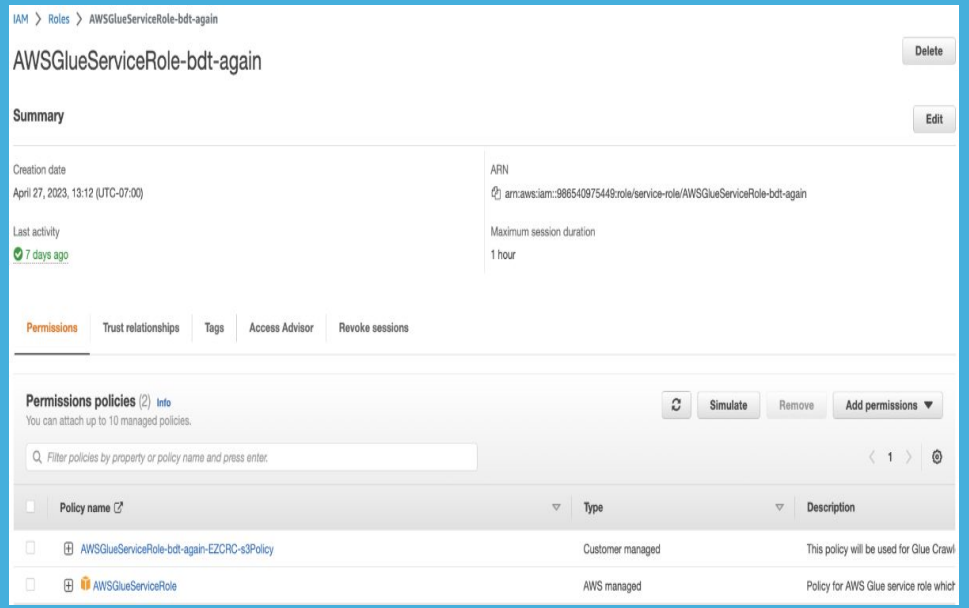
- **Data Cleaning: We cleaned the data to ensure consistency and accuracy by removing duplicates, handling missing values, and correcting data types using PySpark on Google Colab.**

- **ETL Process: We performed Extract, Transform, Load (ETL) operations to prepare the data for analysis. We extracted the data, transformed it to remove inconsistencies, and loaded it into Redshift Data Warehouse using Glue.**

Data Flow using AWS Services



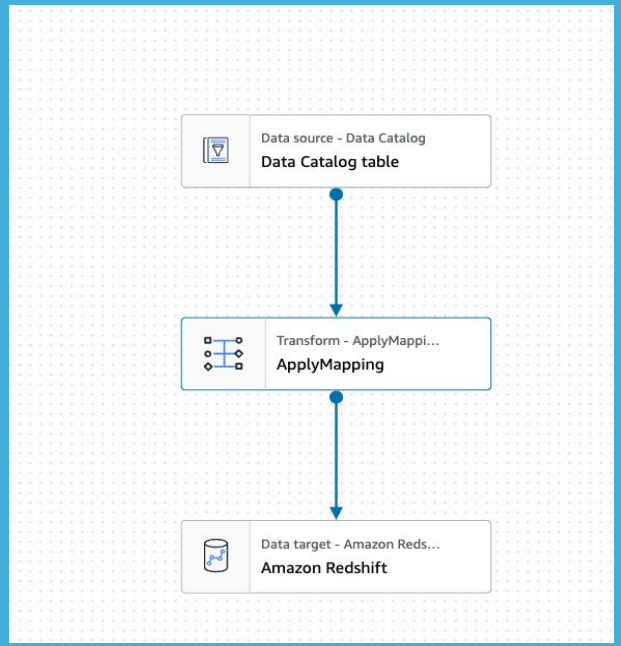
S3 bucket

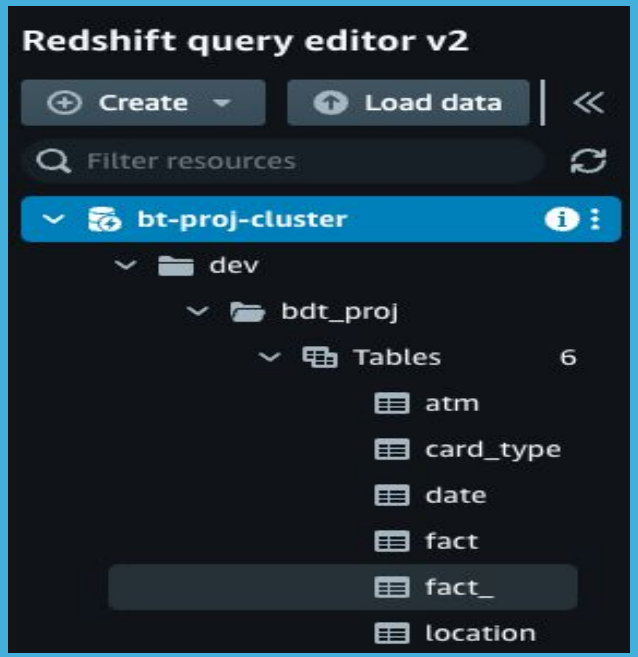


IAM Role

Setting up VPC and Routing table for S3

AWS Glue Job


Data source - Data Catalog
Data Catalog table
Transform - ApplyMappi...
ApplyMapping
Data target - Amazon Reds...
Amazon Redshift


Redshift query editor v2
Create
Load data
Filter resources
bt-proj-cluster
dev
bdt_proj
Tables
6
atm
card_type
date
fact
fact_
location

Data Analytics and Visualization

Data Analytics and Visualization

- Data analytics was performed using Amazon Redshift, a cloud-based data warehouse that allows for fast querying and processing of large datasets.
- Key insights and findings from the data analytics include trends in ATM usage by time of day, day of week, and location, as well as correlations between ATM usage and weather conditions.
- Data visualization techniques used in the project include line charts, bar charts which helped to visually communicate the trends and insights discovered through data analysis.

Top 10 ATMs with the highest percentage of 'inactive' transactions

```sql
--Top 10 ATMs where most transactions are in the 'inactive' state
select a.atm_number, a.atm_manufacturer, l.location, count(a.atm_number) as total_transaction_count, sum(case when
atm_status = 'Inactive' then 1 else 0 end) as inactive_count, round((inactive_count * 100.00 /
total_transaction_count), 4) as inactive_count_percent from bdt_proj.atm a, bdt_proj.location l, bdt_proj.fact_ f
where f.atm_id = a.atm_id and f.weather_loc_id=l.atm_location_id and f.atm_status = 'Inactive' group by a.
atm_number, a.atm_manufacturer, l.location order by inactive_count DESC limit 10;
```

Result 1 (10) | Export | Chart

atm_number	atm_manufacturer	location	total_transaction_count	inactive_count	inactive_count_percent
16	NCR	Skive	44043	44043	100
12	NCR	Østerå Duus	33982	33982	100
2	NCR	Vejgaard	33725	33725	100
88	NCR	Storcenter indg. A	32183	32183	100
30	NCR	Nykøbing Mors	30883	30883	100
52	NCR	Farsø	27361	27361	100
50	NCR	Aarhus	23416	23416	100
29	NCR	Skelagervej 15	20773	20773	100
81	NCR	Spar Købmand Tornhøj	20148	20148	100
102	NCR	Aalborg Storcenter Afd	18297	18297	100

Amount of transactions made annually on weekdays and weekends using an ATM

```
25  --Number of transactions happening on an ATM on weekdays and on weekends throughout the year. Order this by the
    ATM_number, ATM_manufacturer, location, weekend_flag and then total_transaction_count
26  select a.atm_number, a.atm_manufacturer, l.location,case when d.weekday='Thursday' then 1 when d.weekday='Sunday'
    then 1 else 0 end as weekend_flag, count(a.atm_number) as total_transaction_count from bdt_proj.atm a, bdt_proj.
    location l, bdt_proj.fact_ f, bdt_proj.date d where f.atm_id = a.atm_id and f.weather_loc_id=l.atm_location_id and f.
    date_id=d.date_id  group by a.atm_number, a.atm_manufacturer, l.location,weekend_flag order by a.atm_number;
27
```

Result 1 (100) Export Chart

atm_number	atm_manufacturer	location	weekend_flag	total_transaction_count
1	NCR	Næstved	1	11154
1	NCR	Næstved	0	31633
2	NCR	Vejgaard	1	8641
2	NCR	Vejgaard	0	25084
3	NCR	Ikast	0	10019
3	NCR	Ikast	1	3621
4	NCR	Svogerslev	0	25151
4	NCR	Svogerslev	1	8953
5	NCR	Nibe	1	4683
5	NCR	Nibe	0	14058
6	NCR	Fredericia	0	16840

Number of unsuccessful ATM transactions using different card types

```
21  --Number of failed ATM transactions across various card types
22  select c.card_type,count(f.trans_id) as total_transaction_count,sum(case when f.atm_status = 'Inactive' then 1 else
    0 end) as inactive_count, round((inactive_count * 100.00 / total_transaction_count), 4) as inactive_count_percent
    from bdt_proj.fact_ f,bdt_proj.card_type c where f.card_type_id=c.card_type_id group by c.card_type order by
    inactive_count_percent desc;
```

Result 1 (12) | Export | Chart

card_type	total_transaction_count	inactive_count	inactive_count_percent
Mastercard - on-us	458226	86000	18.768
VISA	170828	30713	17.9789
Dankort - on-us	143813	24680	17.1612
CIRRUS	17362	2953	17.0084
Hævekort - on-us	62487	10331	16.533
Dankort	28581	4557	15.9442
MasterCard	400507	63482	15.8504
Visa Dankort - on-us	748805	112972	15.087
Hævekort	8459	1208	14.2806
Visa Dankort	427840	60547	14.1518
VisaPlus	1134	150	13.2275

Most active day (in terms of transactions) in each ATM in Vejgaard location

```
29  --Most active day in each ATMs from location "Vejgaard"
30  select a.atm_number, a.atm_manufacturer, l.location,d.weekday, count(a.atm_number) as total_transaction_count from
    bdt_proj.atm a, bdt_proj.location l, bdt_proj.fact_ f, bdt_proj.date d where f.atm_id = a.atm_id and f.
    weather_loc_id=l.atm_location_id and f.date_id=d.date_id and l.location = 'Vejgaard' group by a.atm_number, a.
    atm_manufacturer, l.location, d.weekday order by d.weekday, total_transaction_count limit 2;
```

Result 1 (2) | Export | Chart

atm_number	atm_manufacturer	location	total_transaction_count	weekday
103	Diebold Nixdorf	Vejgaard	4757	Friday
2	NCR	Vejgaard	6290	Friday

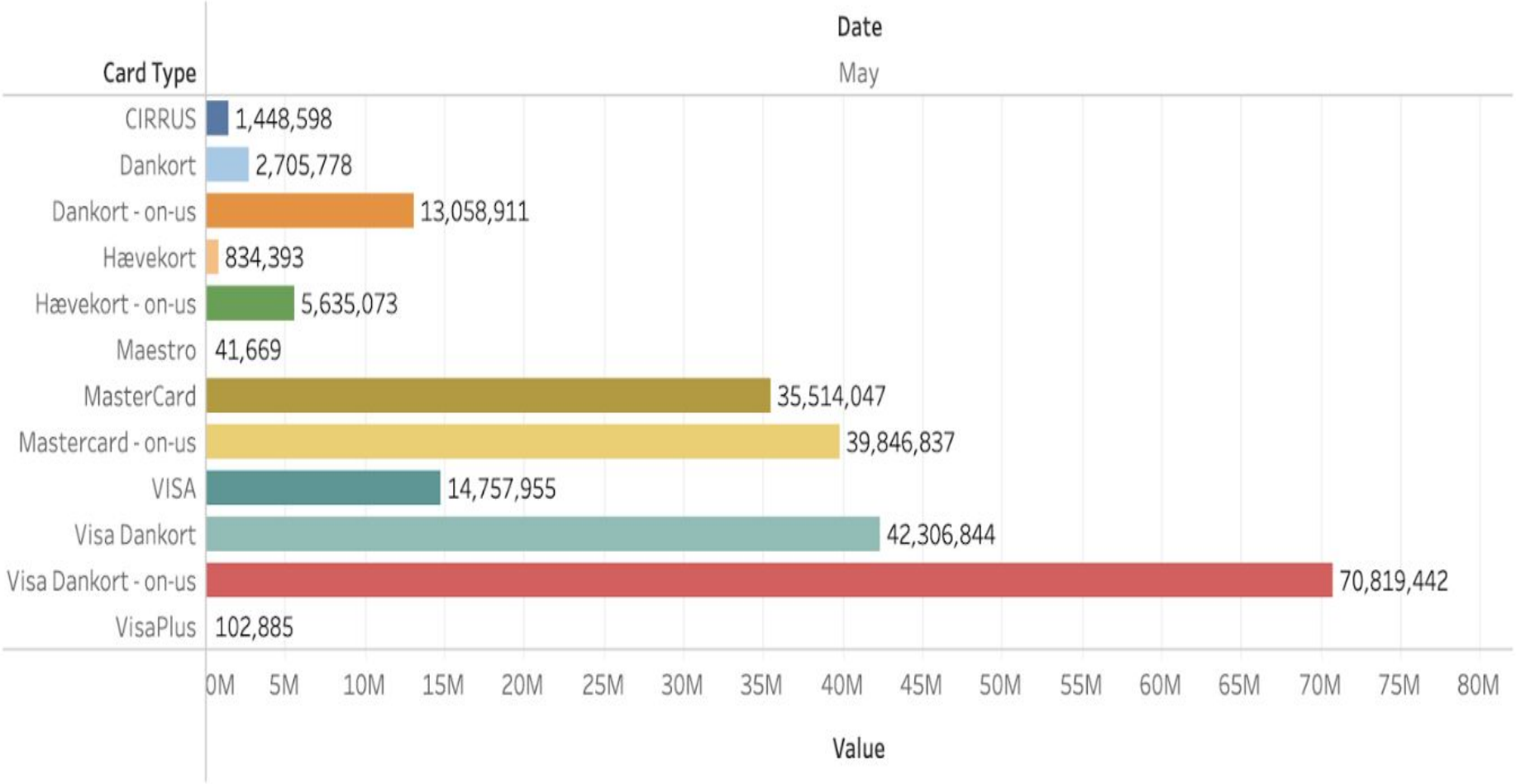
During what time frame most ATM transactions occurred

```
32  -- Amount of transactions at ATMs by each hour sorted in descending order
33  select distinct d.hour, count(*) as no_of_transactions from bdt_proj.date d, bdt_proj.fact_ f where d.date_id = f.
    date_id group by d.hour order by no_of_transactions desc;
```

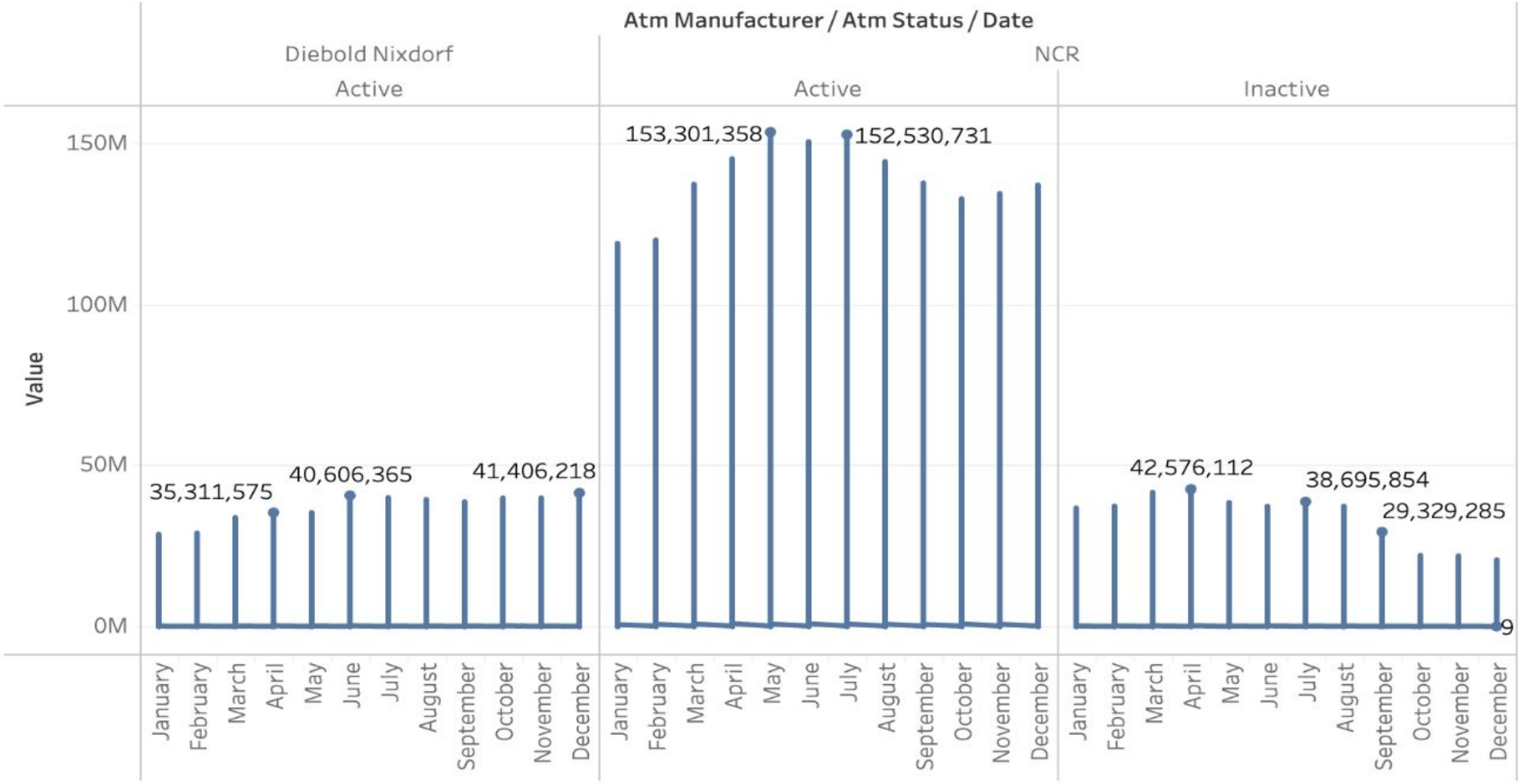
Result 1 (24) | Export | Chart

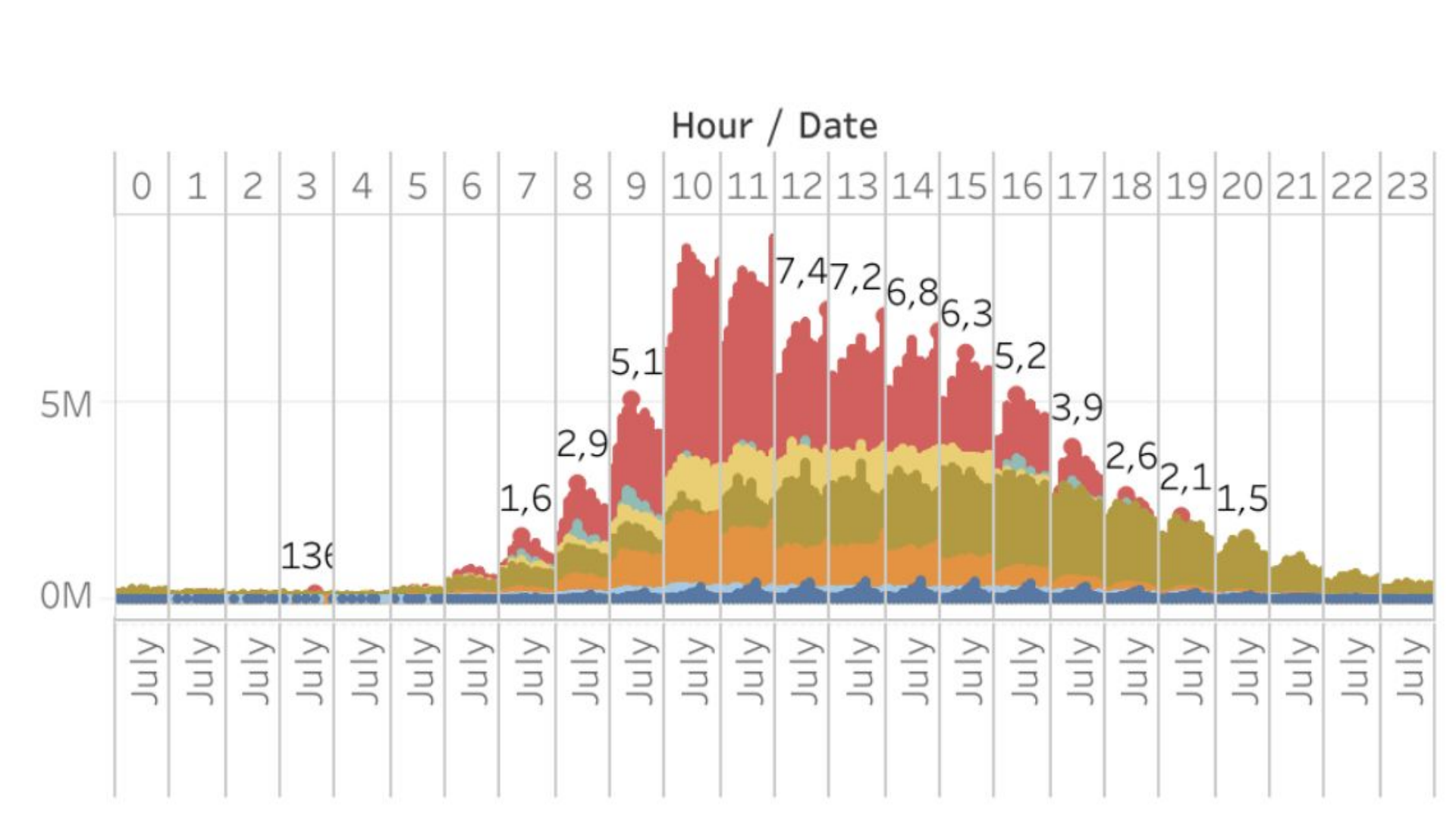
hour	no_of_transactions
11	256542
10	249928
12	237107
13	233322
14	230270
15	223323
16	196004
17	153816
9	150943

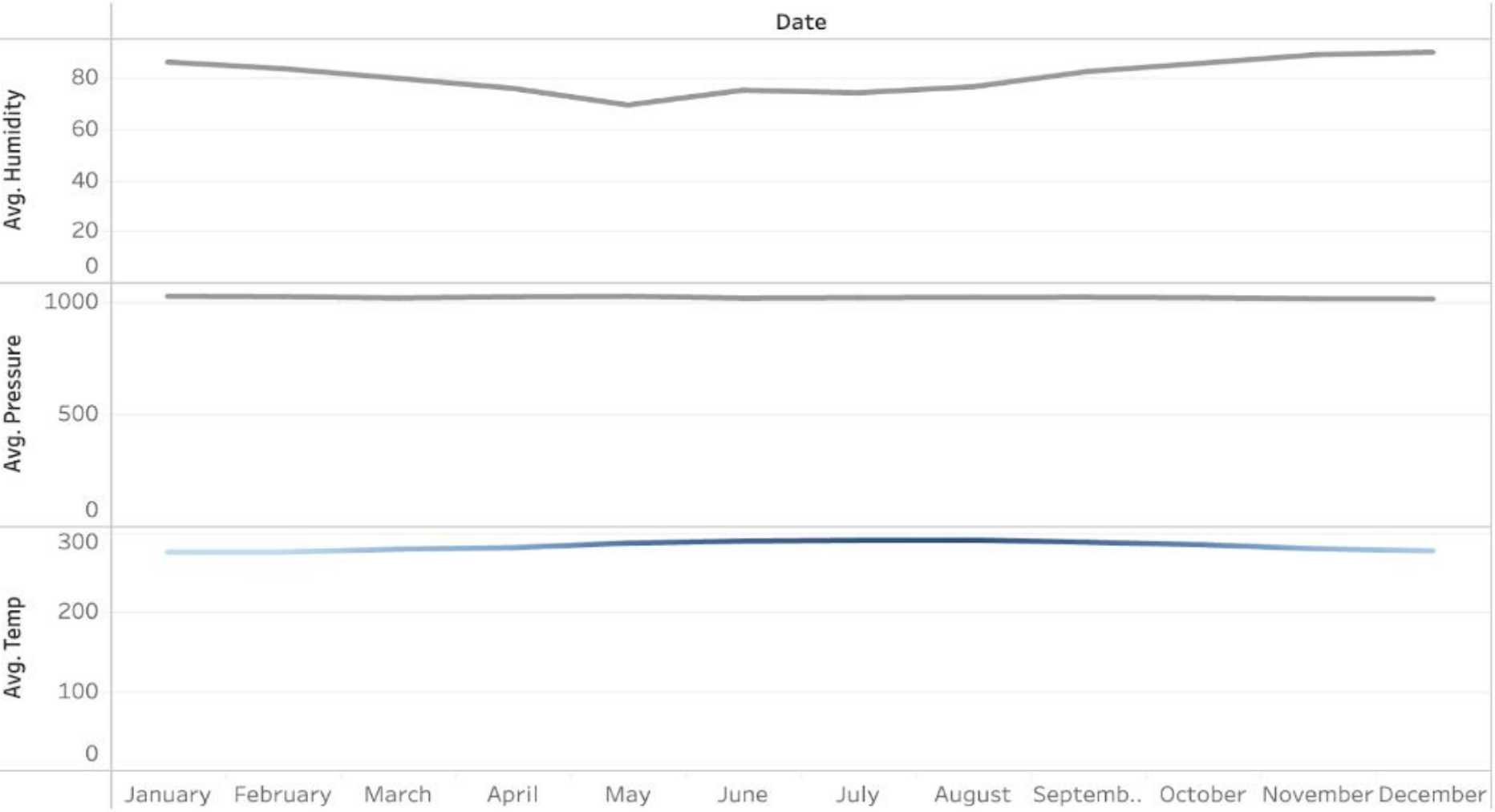
ATM Transactions by Card Type



ATM Transactions by Manufacturer and their Status




Hour / Date
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
5M
0M
1,6
2,9
5,1
7,4
7,2
6,8
6,3
5,2
3,9
2,6
2,1
1,5
July

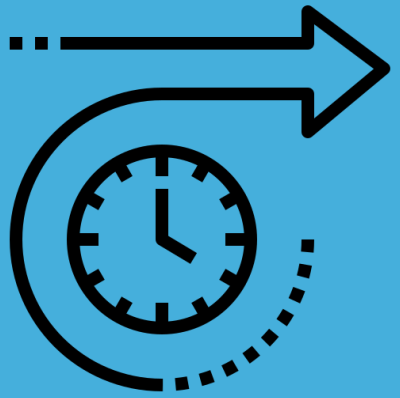

Date
Avg. Humidity
80
60
40
20
0
Avg. Pressure
1000
500
0
Avg. Temp
300
200
100
0
January
February
March
April
May
June
July
August
Septemb..
October
November
December

Conclusion

Conclusion

- **Insights into Customer Behavior:** By analyzing Spar Nord Bank's ATM transactional data, our project provided valuable insights into customer behavior, such as peak transaction times and preferred payment methods. These insights can inform the bank's decision-making processes and help optimize the bank's ATM network.

- **Potential Cost Savings:** Our study revealed that the bank can utilize non-peak time periods to refill the ATMs, potentially saving money on the cost of refilling the machines. This finding demonstrates the potential cost savings that can be achieved through data-driven optimization.



Future Work

Future Work

- **Explore the impact of ATM features on usage:** As mentioned in the previous point, it could be beneficial to investigate how specific features of ATMs (such as deposit capabilities or language options) influence customer usage patterns. This could involve conducting surveys or collecting additional data from ATM transactions to gain more insights into customer preferences.

- **Incorporate demographic data:** To better understand the needs and behaviors of different customer segments, future work could involve collecting and analyzing demographic data on ATM users (such as age, gender, income, etc.). This information could be used to inform decisions about where to locate ATMs, what features to offer, and how to market ATM services to different audiences.

- **Build predictive models:** By leveraging the data gathered from ATM transactions, future work could involve building predictive models to forecast usage patterns, detect anomalies, and identify opportunities for optimization. This could involve using machine learning algorithms to identify patterns and trends in the data, and then using these insights to make informed decisions about ATM operations and strategy.

References

https://dribbble.com/shots/3076212-ATM

2.5M Danish ATM Transactions from 2017. (2019, January 16). Kaggle.https://www.kaggle.com/datasets/sparnord/danish-atm-transactions

https://www.sparnord.com/