IT Portfolio

Dear reader,

This document presents a portfolio created by me, Elizaveta Novikova, using diverse programming languages and software. I hereby confirm that I am the author of this work, and it has not been attributed to anyone else. If you have any questions regarding this document or my qualifications, please feel free to contact me via email at eliza.nkv@gmail.com, by phone at +420735639438, or LinkedIn.

Here is my GitHub page, where you can find all the codes.

Elizaveta Novikova

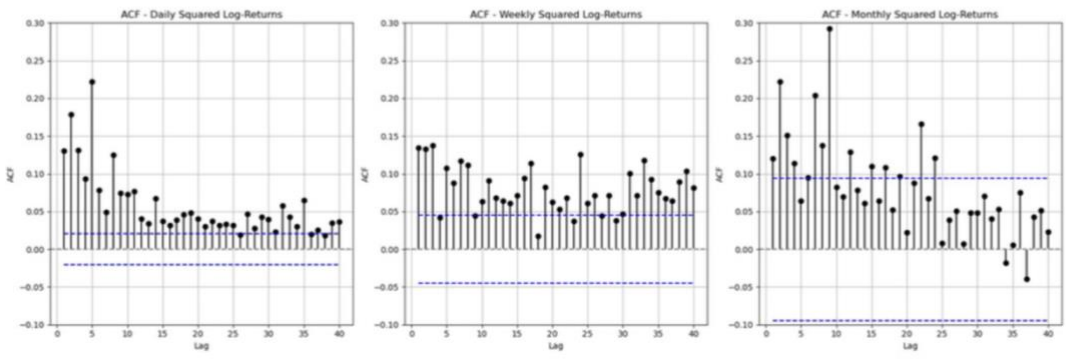
Content

Python analysis

The code then computes the autocorrelation function (ACF) for the squared daily log returns, which represents the autocorrelation of the variance of the stock returns. ACF measures the relationship between the values of a time series at different lags. In this context, it helps analyze the persistence of volatility.

```
In [6]:import yfinance as yf
import matplotlib.pyplot as plt
import numpy as np
from statsmodels.tsa.stattools import acf

# Scarica i dati del MSFT nel periodo specificato
MSFT = yf.download("MSFT", start="1986-12-31", end="2022-12-31")

# Estrai i log-returns giornalieri
log_returns_daily = np.log(MSFT['Adj Close']).diff().dropna()

# Calcola l'autocorrelazione empirica
lags = 40
acf_values_daily = acf(log_returns_daily**2, nlags=lags)

# Calcola le bande di confidenza a 1.96 volte la deviazione standard dell'autocorrelazione
confint = 1.96 / np.sqrt(len(log_returns_daily))
confint_upper = np.full(lags, confint)
confint_lower = -np.full(lags, confint)

# Creazione dei tre grafici affiancati
fig, axs = plt.subplots(1, 3, figsize=(18, 6))

# ACF dei log-returns giornalieri con bande di confidenza
axs[0].stem(np.arange(1, lags + 1), acf_values_daily[1:], linefmt='k-', markerfmt='ko', basefmt='w-')
axs[0].axhline(y=0, color='gray', linestyle='--')
axs[0].plot(np.arange(1, lags + 1), confint_upper, color='blue', linestyle='dashed')
axs[0].plot(np.arange(1, lags + 1), confint_lower, color='blue', linestyle='dashed')
axs[0].set_ylim(-0.1, 0.3)
axs[0].set_title('ACF - Daily Squared Log-Returns')
axs[0].set_xlabel('Lag')
axs[0].set_ylabel('ACF')
axs[0].grid(True)

# ACF dei log-returns settimanali con bande di confidenza
acf_values_weekly = acf(log_returns_weekly**2, nlags=lags)
confint_weekly = 1.96 / np.sqrt(len(log_returns_weekly))
confint_weekly_upper = np.full(lags, confint_weekly)
confint_weekly_lower = -np.full(lags, confint_weekly)

axs[1].stem(np.arange(1, lags + 1), acf_values_weekly[1:], linefmt='k-', markerfmt='ko', basefmt='w-')
axs[1].axhline(y=0, color='gray', linestyle='--')
axs[1].plot(np.arange(1, lags + 1), confint_weekly_upper, color='blue', linestyle='dashed')
axs[1].plot(np.arange(1, lags + 1), confint_weekly_lower, color='blue', linestyle='dashed')
axs[1].set_ylim(-0.1, 0.3)
axs[1].set_title('ACF - Weekly Squared Log-Returns')
axs[1].set_xlabel('Lag')
axs[1].set_ylabel('ACF')
axs[1].grid(True)

# ACF dei log-returns mensili con bande di confidenza
acf_values_monthly = acf(log_returns_monthly**2, nlags=lags)
confint_monthly = 1.96 / np.sqrt(len(log_returns_monthly))
confint_monthly_upper = np.full(lags, confint_monthly)
confint_monthly_lower = -np.full(lags, confint_monthly)

                    axs[2].stem(np.arange(1, lags + 1), acf_values_monthly[1:], linefmt='k-', markerfmt='ko', basefmt='w-')
axs[2].axhline(y=0, color='gray', linestyle='--')
axs[2].plot(np.arange(1, lags + 1), confint_monthly_upper, color='blue', linestyle='dashed')
axs[2].plot(np.arange(1, lags + 1), confint_monthly_lower, color='blue', linestyle='dashed')
axs[2].set_ylim(-0.1, 0.3)
axs[2].set_title('ACF - Monthly Squared Log-Returns')
axs[2].set_xlabel('Lag')
axs[2].set_ylabel('ACF')
axs[2].grid(True)

# Regolazione dello spaziamento tra i grafici
plt.tight_layout()

# Salva il grafico in formato png
plt.savefig('MSFT_rt_SQUAREDrt_d_1986_2022.png', format='png', bbox_inches='tight')

plt.show()
```

```
[**********100%%**********] 1 of 1 completed
```



SQL analysis

A Japanese international university surveyed its students in 2018 and published a study the following year that was approved by several ethical and regulatory boards. Using the data from the csv file, these two analyses showcased that international students have a higher risk of mental health difficulties than the general population, and that social connectedness (belonging to a social group) and acculturative stress (stress associated with joining a new culture) are predictive of depression.

```sql
-- Load data from CSV file into a table
CREATE TABLE students AS
SELECT *
FROM read_csv_auto('students.csv');

-- Create a table for international vs domestic students
SELECT inter_dom,
ROUND(AVG(stay), 2) AS 'Average length of stay (in years)',
ROUND(AVG(todep), 2) AS 'Average score of depression (PHQ-9 test)',
ROUND(AVG(tosc), 2) AS 'Average score of social connectedness (SCS test)',
ROUND(AVG(toas), 2) AS 'Average score of acculturative stress (ASISS test)'
FROM students
WHERE inter_dom IS NOT NULL
GROUP BY inter_dom
```

[2]

	inter_dom	Average length of stay (in years)	Average score of depression (PH...	Average score of social connecte...	Average score of acculturative st...
0	Inter	2.06	8.04	37.42	75.56
1	Dom	2.4	8.61	37.64	62.84

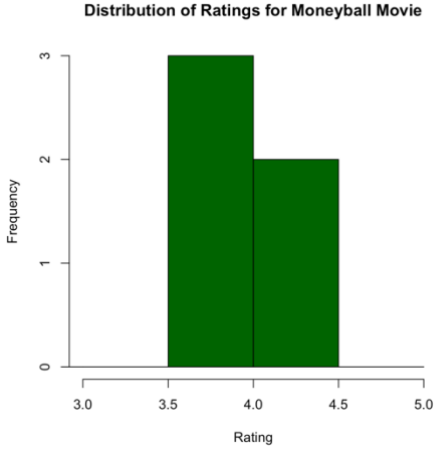
```sql
-- Load data from CSV file into a table
CREATE TABLE students AS
SELECT *
FROM read_csv_auto('students.csv');

-- Create a table to analyze the length of stay as a contributing factor
SELECT stay as 'Length of stay in years',
COUNT(inter_dom) as 'Nº of international students',
ROUND(AVG(todep), 2) AS 'Avg score of depression (PHQ-9 test)',
ROUND(AVG(tosc), 2) AS 'Avg score of social connectedness (SCS test)',
ROUND(AVG(toas), 2) AS 'Avg score of acculturative stress (ASISS test)'
FROM students
WHERE inter_dom = 'Inter'
GROUP BY stay
ORDER BY stay DESC
LIMIT 9;
```

[3]

	Length of stay in years	Nº of international students	Avg score of depression (PHQ-...	Avg score of social connected...	Avg score of acculturative str...
0	10	1	13	32	50
1	8	1	10	44	65
2	7	1	4	48	45
3	6	3	6	38	58.67
4	5	1	0	34	91
5	4	14	8.57	33.93	87.71
6	3	46	9.09	37.13	78
7	2	39	8.28	37.08	77.67
8	1	95	7.48	38.11	72.8

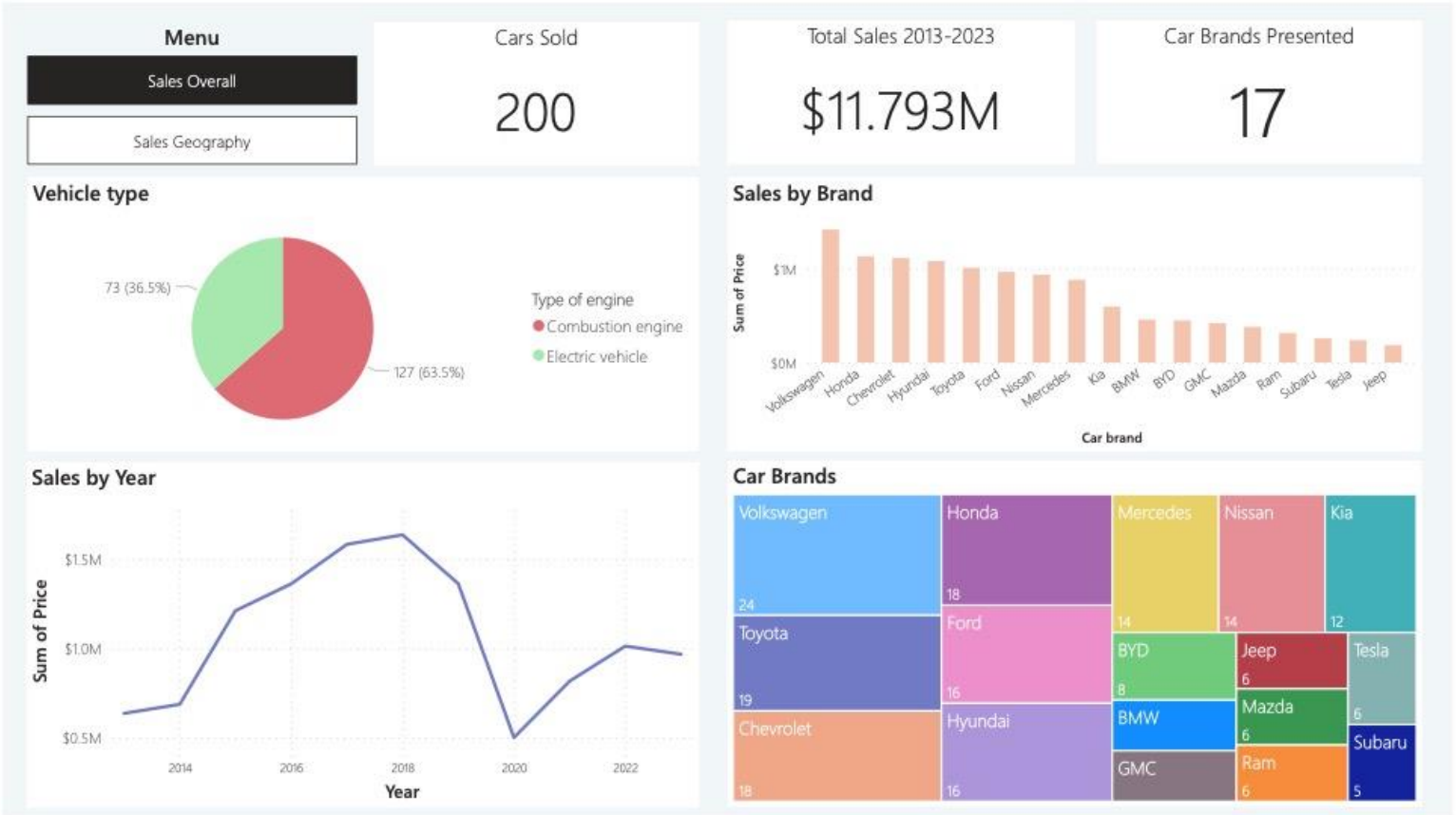
R analysis

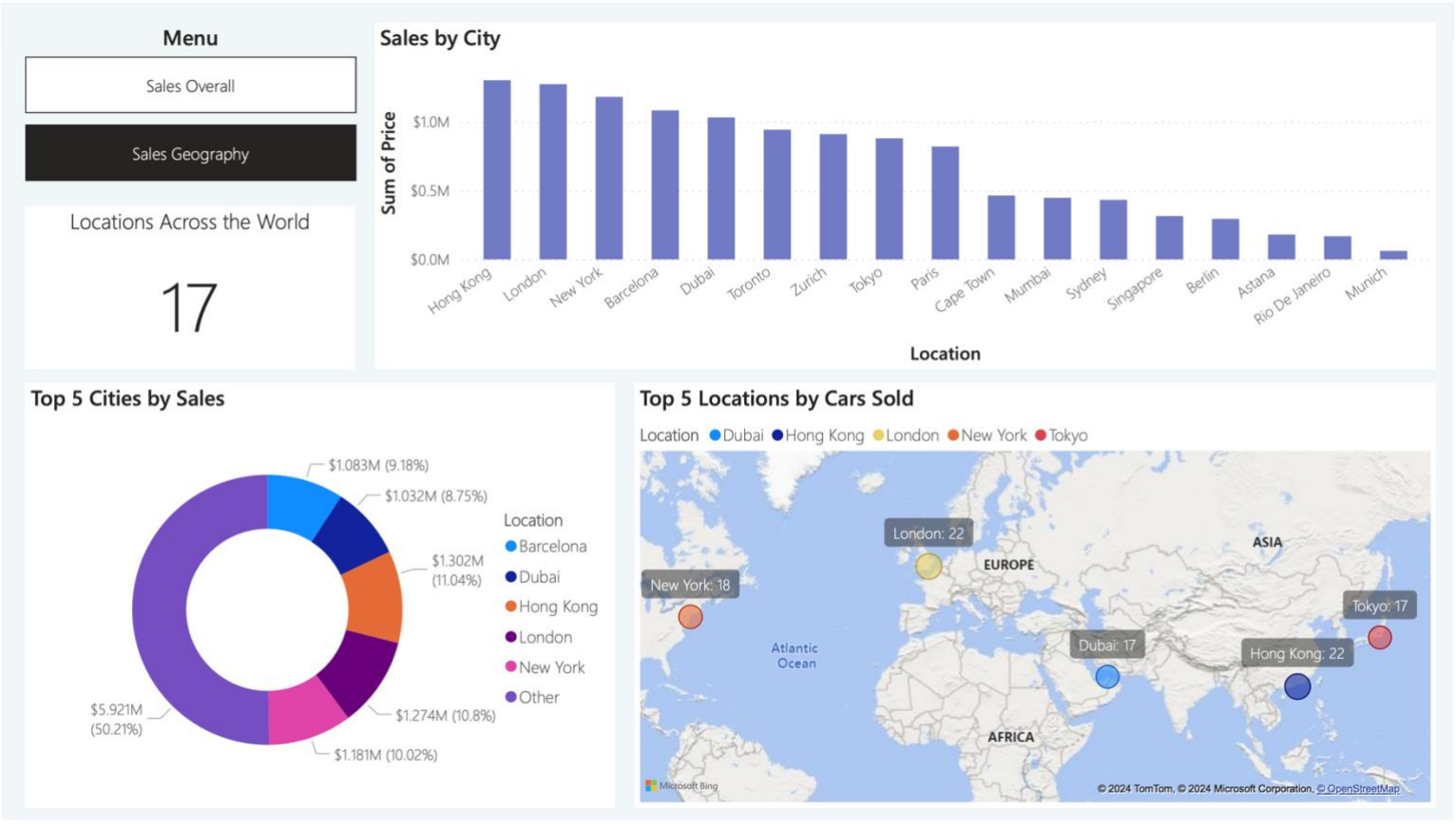
```r
# Movie ratings analysis using R & SQL

#activate SQL
install.packages("RMySQL")
library(RMySQL)

#connect to our database
?dbConnect
our.conn=dbConnect(MySQL(), user="12345", password="12345",
                   dbname="12345", host="12345", port=12345) #due to the security reasons, the data is changed here
our.conn

#find Moneyball movie ID
movies=dbGetQuery(our.conn,'SELECT * FROM movies;')
Moneyball_ID=dbGetQuery(our.conn,'SELECT movieId FROM movies WHERE title="Moneyball";')
Moneyball_ID

#then use movies-ratings table to extract all ratings for Moneyball movie
movies_ratings=dbGetQuery(our.conn,'SELECT * FROM movies_ratings;')

#what is the average rating for this movie?
Moneyball_avg_rating=dbGetQuery(our.conn,'SELECT AVG(rating) FROM movies_ratings WHERE movieId = 89492;')
Moneyball_avg_rating

#how many ratings are there
Moneyball_count_rating=dbGetQuery(our.conn,'SELECT COUNT(rating) FROM movies_ratings WHERE movieId = 89492;')
Moneyball_count_rating

#visualize the distribution of movie ratings
Moneyball_ratings <- dbGetQuery(our.conn, 'SELECT rating FROM movies_ratings WHERE movieId = 89492;')
Moneyball_ratings_df <- data.frame(rating = Moneyball_ratings$rating)
par(mfrow=c(1,1))
hist(Moneyball_ratings_df$rating,
     main="Distribution of Ratings for Moneyball Movie",
     xlab='Rating',col='darkgreen',xlim=c(3, 5),breaks=seq(0, 5, by=0.5))

#is this a good score? find out what is average,
movies_ratings_avg=dbGetQuery(our.conn,'SELECT AVG(rating) FROM movies_ratings')
movies_ratings_avg

#max,
movies_ratings_max=dbGetQuery(our.conn,'SELECT MAX(rating) FROM movies_ratings')
movies_ratings_max

#and min score on table movies_ratings
movies_ratings_min=dbGetQuery(our.conn,'SELECT MIN(rating) FROM movies_ratings')
movies_ratings_min

print('Results: Monetball movie has a rating above average, therefore its score is fairly good')

#The output of the hist() function:
```



Power BI project

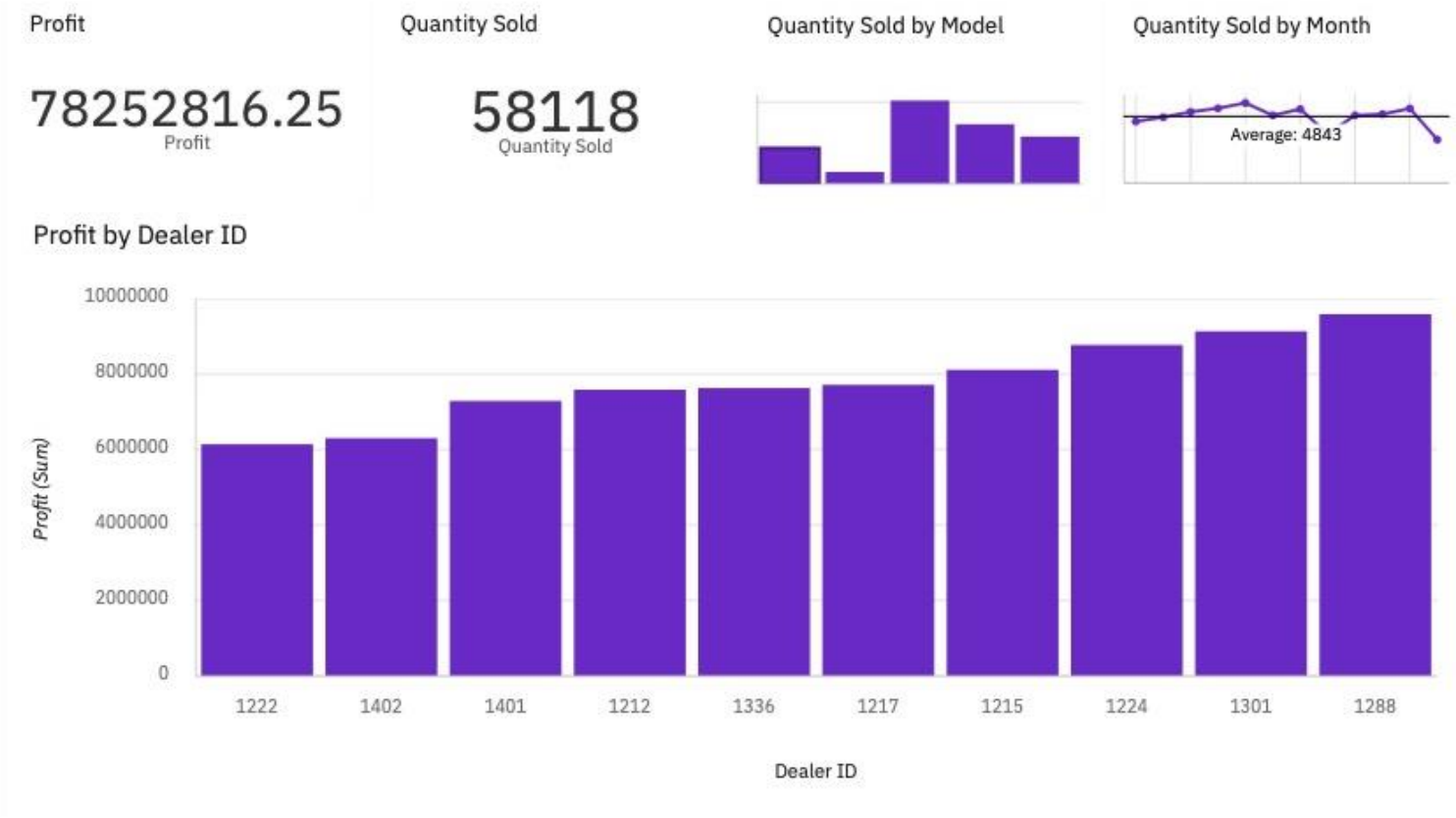
A random data set for an imaginary Vehicle Dealers company was made up by the author and stored in an Excel file.

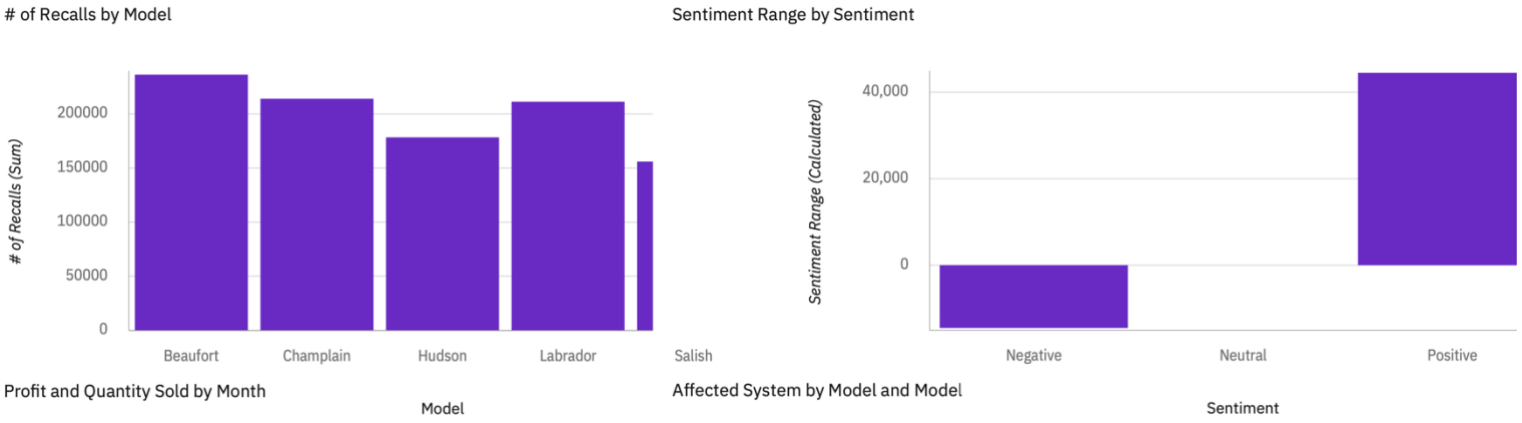


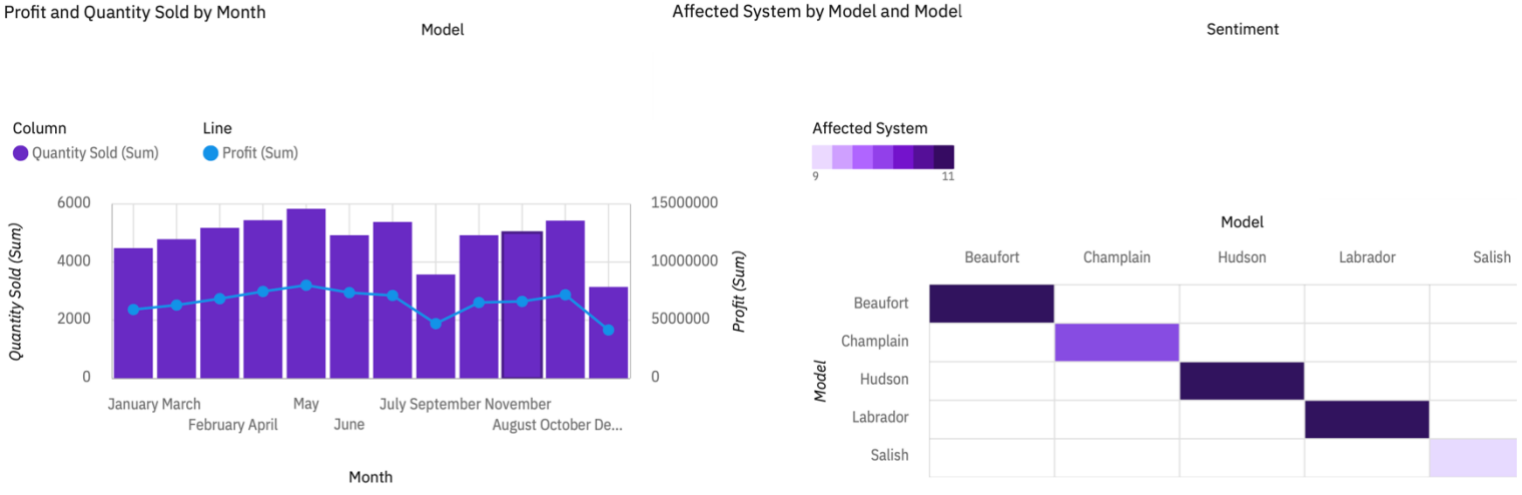


Cognos IBM project

A random data set for an imaginary Vehicle Dealers company was made up by the author and stored in an Excel file.







HTML & CSS

My “business card” website done by using HTML and CSS. ⇒ The URL ⇐

Elizaveta Novikova

Hi there! I am Liza, a graduate of Prague University of Economics and Business and EDHEC with a strong academic record and a passion for Data Science and Risk Management. I am motivated and adaptable, eager to contribute to an international environment.

Education & Experience | Skills & Interests | Portfolio & Resume | Contact

Education 🎓

University	Degree
EDHEC	Erasmus+, BSc in Finance
University of Economics and Business (VŠE)	Bachelor of Business Administration
Charles University	Foundation year, Economics in Czech

Experience 👩‍💻

Employer	Position	Dates
Clearstream (Deutsche Börse Group)	Business Intelligence Specialist – Risk Management team	Jan 2024 – Present
Deutsche Börse	Service Desk Analyst	Jan 2023 – Sep 2023
Yandex	Data Analyst Intern	May 2022 – Sep 2022

Skills and interests 🎨

Programming languages:

Python SQL R DAX VBA HTML CSS JavaScript

Software:

Power BI Tableau Cognos Jira SAP products Power Apps Adobe Photoshop

Languages:

English fluent Russian fluent Belarusian fluent Czech B2 German B1 French A2

Interests:

Art Classical Music Learning Languages Sports (Gym, Football, Hiking, Chess)

Documents 📋

Download My Portfolio
Download My Resume
My LinkedIn
My GitHub

Contact Me 📞

Interested in working together? Feel free to reach out!

Your Name

Your Email

Your Message

Send Message