PROJECT PROPOSAL

Saudi Stock Exchange

Problems

Using data science in the stock market is not new, but that doesn't apply for Saudi Stock Exchange (Tadawul), It needs to be explored and studied deeply, so we can cluster companies based on its behavior during the good and bad days. Also we can identify the days with a very large number of trades and try to understand the reason behind it. Finally we can predict the stocks prices.

Column

Column	describtion	Data type
symbol	The symbol or the reference number of the company	int
name	Name of the company	object
trading_name	The trading name of the company	object
sectoer	The sector in which the company operates	object
date	The date of the stock price	object
open	The opening price	int
high	The highest price of the stock at that day	int
low	The lowest price of the stock at that day	int
close	The closing price	int
change	The change in price from the last day	int
perc_Change	The percentage of the change	int
volume_traded	The volume of the trades for the day	int
value_traded	The value of the trades for the day	int
no_trades	The number of trades for the day	int

View data

	symbol	name	trading_name	sectoer	date	open	high	low	close	change	perc_Change	volume_traded	value_traded	no_trades
0	2030	Saudi Arabia Refineries Co.	SARCO	Energy	2020-03-05	35.55	35.85	34.90	34.90	-0.40	-1.13	436609.0	15399073.50	804.0
1	2030	Saudi Arabia Refineries Co.	SARCO	Energy	2020-03-04	34.70	35.65	34.50	35.30	0.25	0.71	737624.0	25981391.35	1268.0
2	2030	Saudi Arabia Refineries Co.	SARCO	Energy	2020-03-03	34.70	35.15	34.70	35.05	1.05	3.09	489831.0	17116413.40	854.0
3	2030	Saudi Arabia Refineries Co.	SARCO	Energy	2020-03-02	35.20	35.65	34.00	34.00	-0.55	-1.59	736157.0	25858700.60	1242.0
4	2030	Saudi Arabia Refineries Co.	SARCO	Energy	2020-03-01	35.35	35.60	34.25	34.55	-2.05	-5.60	738685.0	25747967.55	1625.0

Info data

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 593819 entries, 0 to 593818
Data columns (total 14 columns):
 #   Column         Non-Null Count    Dtype
---  ------         --------------    -----
 0   symbol         593819 non-null   int64
 1   name           593819 non-null   object
 2   trading_name   593819 non-null   object
 3   sectoer        593819 non-null   object
 4   date           593819 non-null   object
 5   open           587364 non-null   float64
 6   high           587122 non-null   float64
 7   low            587122 non-null   float64
 8   close          593819 non-null   float64
 9   change         593819 non-null   float64
 10  perc_Change    593819 non-null   float64
 11  volume_traded  593819 non-null   float64
 12  value_traded   593819 non-null   float64
 13  no_trades      586128 non-null   float64
dtypes: float64(9), int64(1), object(4)
memory usage: 63.4+ MB
```

T5 Bootcamp

Tools

Tool	Describtion
Python programing	commonly used for developing websites and software, task automation, data analysis, and data visualization
Jupyter Notebook	one of the ideal tools to help you to gain the data science skills you need
NumPy	Python library used for working with arrays
Pandas	Open source Python package that is most widely used for data science
Matplotlib	cross-platform, data visualization and graphical plotting library for Python and its numerical extension NumPy
PowerPoint	The program uses slides to convey information rich in multimedia
Link use	https://www.kaggle.com/salwaalzahrani/saudi-stock-exchange-tadawul

T5 Bootcamp

questions

1. what are the top companies in the Saudi Stock Market(Tadawul)?
2. what is the highest and lowest performance day in the Saudi Stock Market reached (Tadawul)?
3. what are the lowest companies in the Saudi Stock Market (Tadawul)?
4. What are the biggest exchange prices and what companies?
5. In which companies were the biggest exchange?