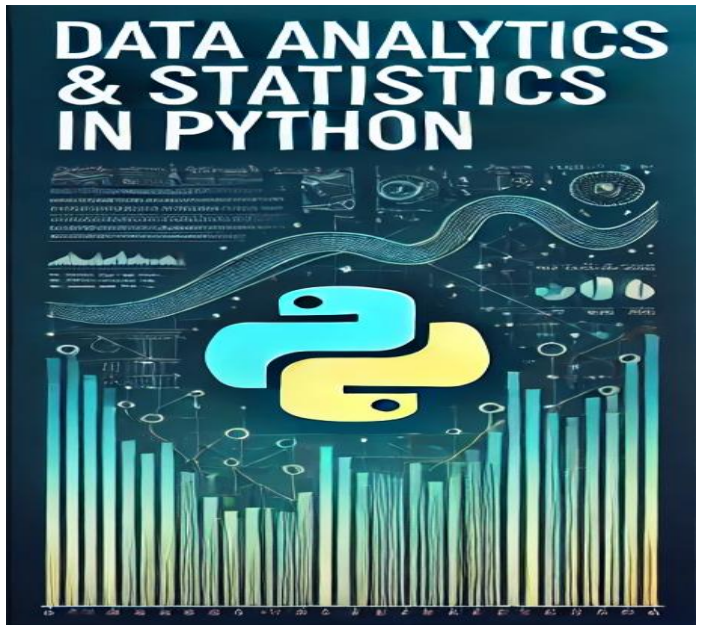
Data Analytics & Statistics in Python
Session 7: Cryptocurrency Analysis Mini-Project



Metropolia
University of Applied Sciences

Learning data-driven decision-making with Python

Instructor: Hamed Ahmadinia, Ph.D.

Email: hamed.ahmadinia@metropolia.fi

Concepts of Today

Session Agenda:

- Cryptocurrency Mini-project Overview
- Jupyter Notebook Walkthrough
- Descriptive Stats, Visualisation & Hypothesis Testing
- Predictive Insights & Token Recommendation
- Kahoot Quiz

Load the Dataset and Explore Basic Information

Load	Load the cryptocurrency dataset (2015–2025)
Inspect	Inspect structure using .head() and .info()
Check	Check for missing values and data types
Initial	Initial shape and data cleaning steps

```python
# Step 1: Loading Dataset
file_path = "crypto_market_data_2018_2024.csv"  # Define the file path to the dataset
df = pd.read_csv(file_path)  # Load the dataset into a DataFrame

df['dates'] = pd.to_datetime(df['dates'])  # Convert the 'dates' column to datetime format for time-series analysis

df.head()  # Display the first 5 rows of the DataFrame to inspect the data
```

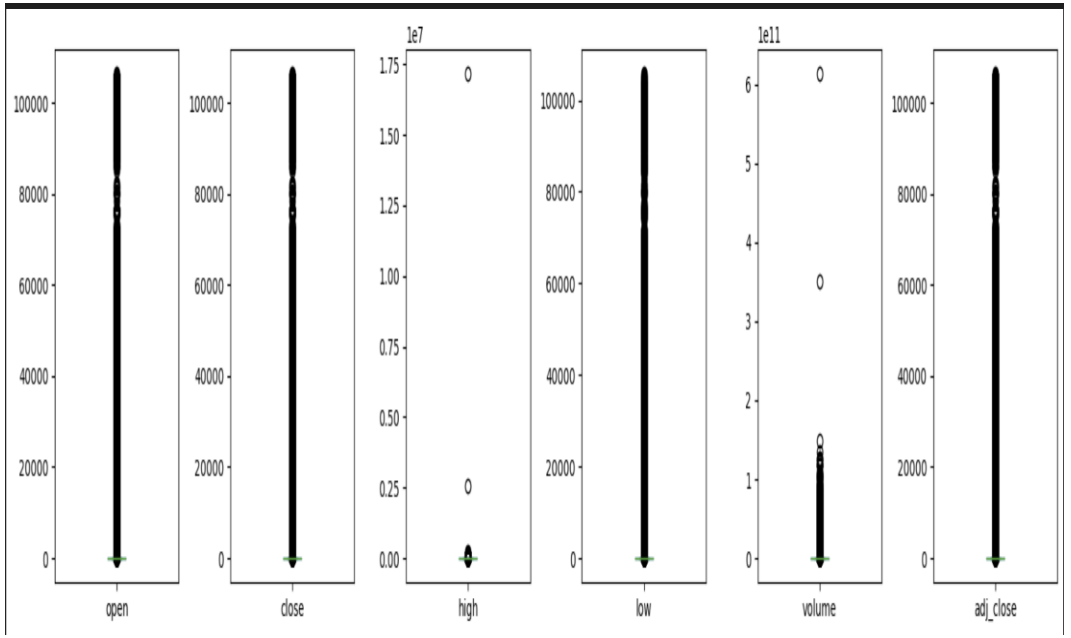
	dates	symbol	open	close	high	low	volume	adj_close
0	2018-01-15	TEL-USD	0.004678	0.006031	0.007141	0.004678	842193.0	0.006031
1	2018-01-16	TEL-USD	0.006056	0.004935	0.006077	0.004112	573317.0	0.004935
2	2018-01-17	TEL-USD	0.004989	0.004539	0.005347	0.003257	477139.0	0.004539
3	2018-01-18	TEL-USD	0.004591	0.007200	0.008505	0.004443	15296600.0	0.007200
4	2018-01-19	TEL-USD	0.007133	0.008325	0.008325	0.006071	15603100.0	0.008325

Perform a descriptive summary of the dataset

Basic metrics: mean, median, standard deviation of price, volume, and market cap

Identify trends by year and by cryptocurrency token

Detect unusual values or outliers using .describe() and visual tools (boxplots, z-scores)

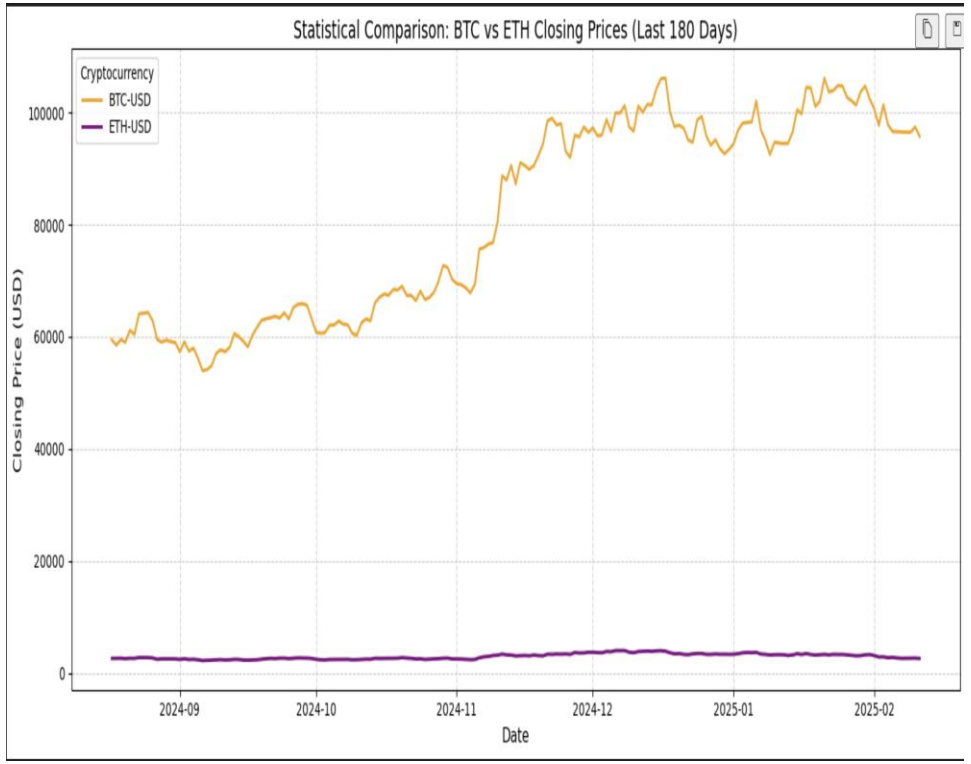


Perform a time-based analysis

GROUP PRICE, VOLUME, AND MARKET CAP DATA BY MONTH AND YEAR

VISUALIZE LONG-TERM TRENDS ACROSS 2015–2025

IDENTIFY **MAJOR SHIFTS** IN TOKEN PERFORMANCE OVER TIME (E.G., BULL/BEAR PHASES)



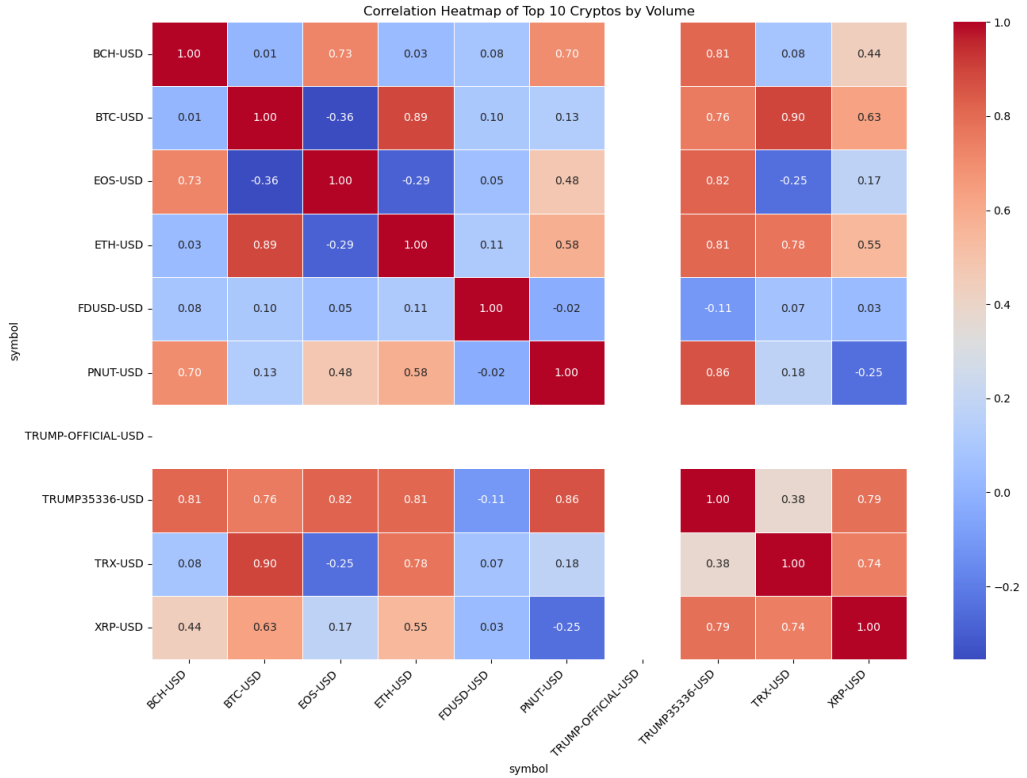
Visualize Data

Histograms: Token popularity and distribution

Boxplots: Detect outliers in price and volume

Line charts: Explore market trends from 2015 to 2025

Heatmaps: Visualize correlations among key features



Predictive Analysis (Optional)

Metropolia
University of Applied Sciences

- **Identify variables affecting token price** trends
- Use **regression or time-series analysis** to model price movement
- **Make a final recommendation**: Which token(s) might be profitable to invest in?

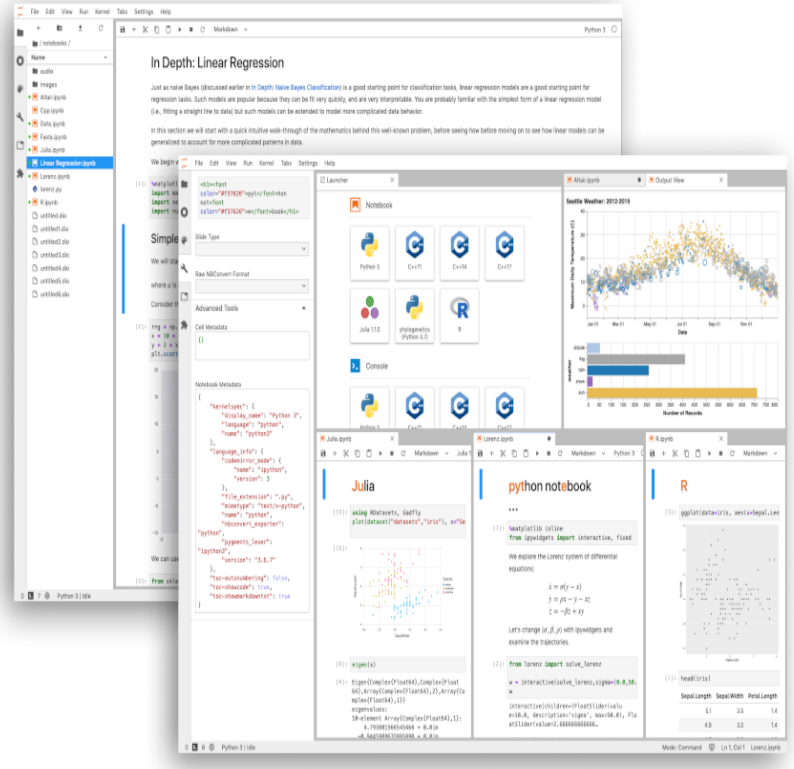
Top 10 Recommended Cryptocurrencies by Model Performance (Positive Growth Only):

	symbol	current_price	predicted_price	growth_rate	MAE	\
333	SUSDE-USD	1.151583	1.167743	0.014032	0.003069	
306	UNFI-USD	0.336828	0.475860	0.412766	0.843185	
11	ZERO31076-USD	0.000110	0.000133	0.205024	0.000034	
140	DUKO-USD	0.000208	0.000326	0.565166	0.000208	
7	BLAST28480-USD	0.004059	0.004739	0.167299	0.001509	
216	HEZ-USD	3.605058	3.822553	0.060330	0.122043	
262	MERL-USD	0.096103	0.125313	0.303935	0.053729	
381	TAIKO-USD	1.049121	1.238912	0.180905	0.167899	
210	JITOSOL-USD	225.270432	240.624095	0.068157	30.517831	
54	ETH-USD	2595.514893	3084.196290	0.188279	508.119174	

Notebook Review

Notebook Walk-through

- **Project Title**: Cryptocurrency Historical Data Analysis
- **Dataset**: Crypto historical data (2015–2025)
- **Goals**:
  - Clean and preprocess data
  - Compute descriptive statistics and visualize trends
  - Conduct hypothesis testing on market behavior
  - Develop predictive models for token price movement
  - Deliver actionable recommendations for potential profitable investments



Kahoot Quiz Time!

Kahoot!

Let’s Test Our Knowledge!

Reference

- Vohra, M., & Patil, B. (2021). A Walk Through the World of Data Analytics. , 19-27. https://doi.org/10.4018/978-1-7998-3053-5.ch002.
- VanderPlas, J. (2016). Python data science handbook: Essential tools for working with data. O'Reilly Media. Available at https://jakevdp.github.io/PythonDataScienceHandbook/
- Severance, C. (2016). Python for everybody: Exploring data using Python 3. Charles Severance. Available at https://www.py4e.com/html3/
- McKinney, W. (2017). *Python for data analysis: Data wrangling with pandas, NumPy, and Jupyter*. O'Reilly Media.