|  |  |
| --- | --- |
| **Practicum Case** |  |
| COMP6140 | COMP6140001 | COMP6140049  Data Mining |
| **Computer Science** | **O191-COMP6140-NP01-03** |
| ***Valid on*** *Even Semester Year 2019/2020* | **Revision 00** |

**Learning Outcome**

* LO1 – explain concept of data and data preprocessing

**Topic**

* Session 03 – Data Preprocessing Using R

## Sub Topics

* Data Preprocessing
* Aggregation

## Soal

*Case*

**NPEX Cryptocurrency Market**

**NPEX** is a grand cryptocurrency exchange market which mainly focuses exchange of **IDR** (Rupiah) with various major cryptocurrencies such as **BTC** (Bitcoin), **ETH** (Ethereum), and **XRP** (Ripple). To maintain its exponential growth, NPEX is currently adapting the use of **R** to analyze its market data. As a data scientist of NPEX company, you are asked to **analyze the market data** according to the following tasks.

1. **Read** the **CSV** file and **divide** the data based on their **cryptocurrency**.
2. Since the data is shuffled, **order** the data based on **date** and **time** in **ascending** order.
3. **Calculate** **Open** and **Close** price of **ETH** (Ethereum) on **16 August**.

**Notes:** open price means the first price of the market when it opened on the day, and close price means the last price in the end of the day.



***Figure 1.1. Ethereum Open Price***



***Figure 1.2. Ethereum Close Price***

1. **Calculate High** and **Low** price of **BTC** (Bitcoin) on **15 August**.

Notes: high price means the highest price of the market on the day, and low price means the lowest price of the market.

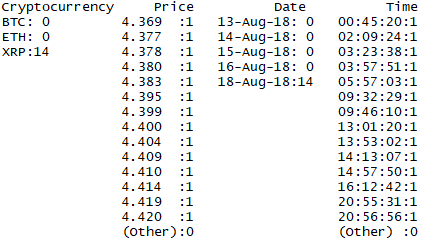


***Figure 2.1. Bitcoin Low Price***



***Figure 2.2. Bitcoin High Price***

1. **Summarize** the price of **XRP** (Ripple) on **18 August** and calculate its **Mean**, **Median**, and **Standard Deviation**.



***Figure 3.1. Ripple Summary***



***Figure 3.2. Ripple Average Price***



***Figure 3.3. Ripple Median Price***



***Figure 3.4. Standard Deviation for Ripple Price***