DATA2001

Data Science, Big Data
and Data Variety

THE UNIVERSITY OF SYDNEY

Outline

- Dataset Introduction – 10 mins
- Point-Based Representation Queries – 30 mins
- Sequence-Based Representation Queries – 40 mins
- Inserting Data and Visualisations – 15 mins
- Group Assignment Reminders – 5 mins

Dataset

O: 33645.54
C: 33750.5
H: 33757.99
L: 33638.85

33270.66

14:26   15:08   15:50   16:31   17:13   17:55   18:36   19:18   04.08.2021 - 20:00   20:41   21:23   22:05

Image source: Better Programming

Image source: Reuters

Financial information is perhaps one of the popular associations made with the world of data, given that being able to well analyse and predict this data could facilitate access to substantial wealth.

The rise of **cryptocurrencies** has particularly proven fascinating over the past decade. Today we'll be investigating the trading price of **Bitcoin** over the past few years as our source of time series data.

Time Series Storage

Today's tutorial revolves around two ways to represent this information:

date	high	low
11-Mar-23 00:00:00	39441.61	39154.63
11-Mar-23 01:00:00	39342.98	38600.00
11-Mar-23 02:00:00	38757.44	38254.50

Point-based representation

date	high	low
11-Mar-23	[39441.61, 39342.98, 38757.44]	[39154.63, 38600.00, 38254.50]

Sequence-based representation

Tasks

To be attempted first in the point-based representation, then in the sequence-based!

Image source: NewsBTC

a) What is the <u>lowest and highest close price</u> that Bitcoin reached in 2023?

b) What was the <u>average midnight open price</u> of Bitcoin over the last month (April 2023), to 2 decimal places?

c) What was the <u>average hourly volume</u> (in $US) traded in July of each year?

Group Assignment

We'll take a few minutes now to talk through the rubric on Canvas. A few other pointers with the group assignment:

- Please double check your **group membership on Canvas** and ensure it correctly reflects your team members. If there are any issues, please post privately on Ed or contact Jarrod (jarrod.jones@sydney.edu.au) ASAP.

- For those without a group, there is now an **Ed megathread** to pair off those remaining. Find a teammate from your tutorial there. Anyone left without a group on Canvas by Sunday will be randomly assigned, but please do your best to organise your own group where possible.

- Do be reminded that we take **plagiarism** very seriously in this unit – ensure the content produced by you (and your teammates!) is your own work.

- Otherwise good luck with it, and direct any **questions** you have to Ed.

Enjoy your week!