SI100B: Real-world Social Exploration Based on Python Data Processing

Kyran Fang

Agenda

- Background
- Introduction of Project
- Grading Criteria

Agenda

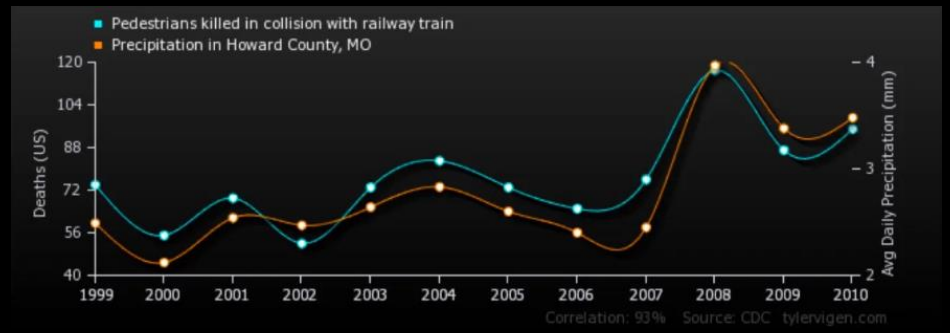
- Background
- Introduction of Project
- Grading Criteria

Background:
Inferring from data is a common thing!

Intricate and varied data are appearing and disappearing every day - or rather, are simply being overwritten by new data. We can use those data to understand the world better.
However, misunderstanding of data is growing by leaps and bounds.

Therefore, mastering data literacy is essential: This requires the ability to not only carry out statistical analysis on real-world problems, but also to understand and critique any conclusion drawn by others on the basis of statistics.

Correlation BUT NOT Causality



	Percentage Still Alive After Five Years		
	All Subjects	*Women*	*Men*
Broccoli lovers (eat it at least once a week)	99	99	99
Broccoli tolerators (eat it about once a month)	98	99	97
Broccoli shunners (wouldn't touch it with a ten-foot pole)	97	99	95

Kyran_

Agenda

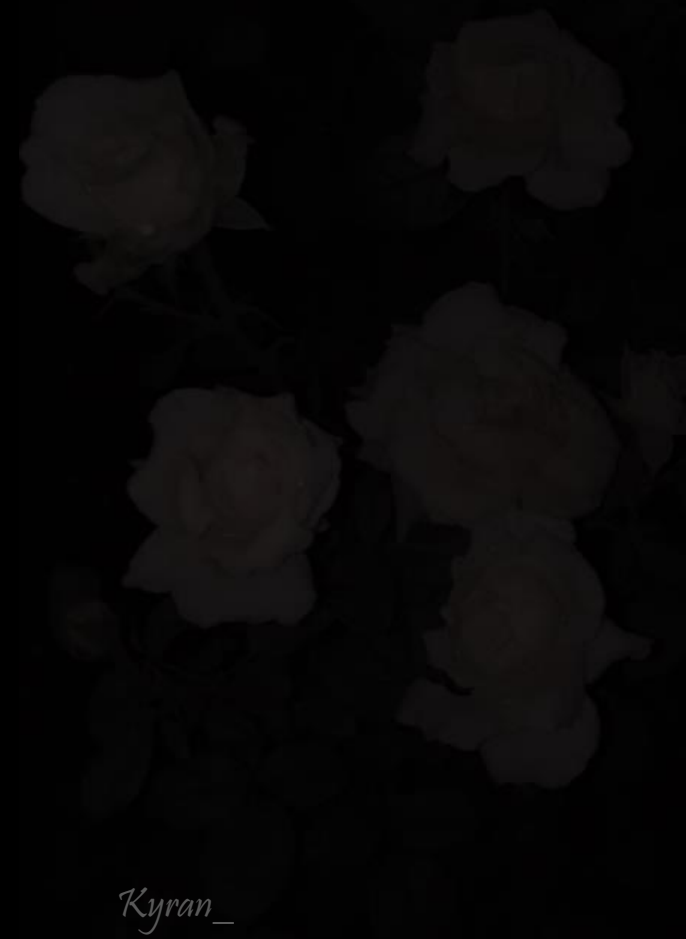
- Background
- Introduction of Project
- Grading Criteria

You will learn:

- Python Data Collection and Basic Data Handling Methods
- Causality and Correlation Analysis
- Pattern Recognition and Deep Learning in Data Science
- Data Visualization

Agenda

- Background
- Introduction of Project
- Grading Criteria



Kyran_

Grading Criteria

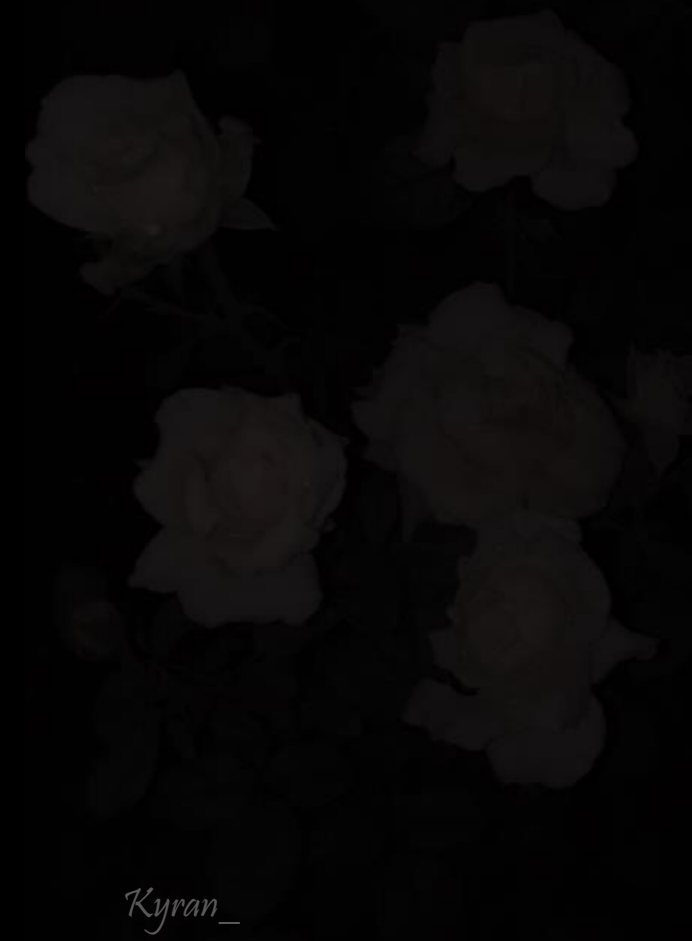
Assessment is divided into two main parts:
Exercise (30%) and Task (70%)

Exercise is a small weekly assignment mainly designed to help you get up to speed quickly and learn how to work with data.

Task is your final challenge, which means that you need to pick a topic to explore and then find the data yourself and give certain conclusions through analysis.

Thanks.

Preference:



Kyran_