

**\* Agenda \***

- \* Types of Data visualization
- \* Univariate Data visualization
  - ✓ Categorical
    - ✓ Bar chart
    - ✓ Count plot
  - ✓ Continuous / Numerical
    - ✓ Histogram
    - ✓ KDE
    - ✓ Box and Whiskers plot
- \* Bivariate Data visualization
  - ✓ Continuous - continuous
    - ✓ Line plot
    - ✓ styling

Module      Numpy

\*\* Assignments

(Q) How to choose the right plot?

(i) How many variables/features are involved?  
 (ii) Whether the variable is numerical / categorical?

↓

|               |   |                     |   |
|---------------|---|---------------------|---|
| * 1-variable  | ⇒ | Univariate analysis | } |
| * 2-variable  | ⇒ | Bivariate "         |   |
| * 2+ variable | ⇒ | Multivariate "      |   |

Page 2 of 14

① Univariate analysis

- ✓ 1. Numerical
- ✓ 2. Categorical

② Bi-variate Analysis

|         |  |
|---------|--|
| ① N - N | N - N - N<br>C - C - C<br>N - C - C<br>C - N - N |
| ② C - C |  |
| ③ N - C |  |

~~---~~

③ Multi variable analysis

(3)

Page 3 of 14

Uni variate Data visualization - Categorical Data

Q what kind of questions we may want to ask for a categorical data?

i) What is the distribution / frequency of the data across diff categories

ii) what proportion does a particular category

Count/proportion

Page 4 of 14

Page 4 of 14

Univariate Data Visualization → Numerical Data

- (i) Some special numbers → min, max, median...
- (ii) Are there any outliers?
- (iii) How is the data distributed
- (iv) How much percentage of the data is above/below a certain number

Distribution

Statistics

Snapshot saved.

The image shows a screenshot of a Chrome browser window. The title bar includes standard menu items like File, Edit, View, History, Bookmarks, Profiles, Tab, Window, Help, and system icons. The date and time 'Tue 20 Jun 11:28 PM' are visible in the top right. The main content area contains handwritten text and diagrams. At the top, 'Univariate Data Visualization' is written in blue, with an arrow pointing to 'Numerical Data' in green. Below this, four questions are listed in a bulleted format. To the left of the first question, 'Distribution' is written in pink. To the right of the fourth question, 'Statistics' is circled in pink, with a green arrow pointing down to a pink-bordered box containing the text 'Snapshot saved.'. The bottom of the slide features a toolbar with various icons.







