



Data Analysis Fundamentals

Lecture 5: Excel Refresher

Vikas Kumar
Indian Institute of Technology Bombay

09th September 2023

Outline

- Data Visualization
- Visualization in Excel
- Practice Problem

Data Visualization

A picture is worth a thousand words

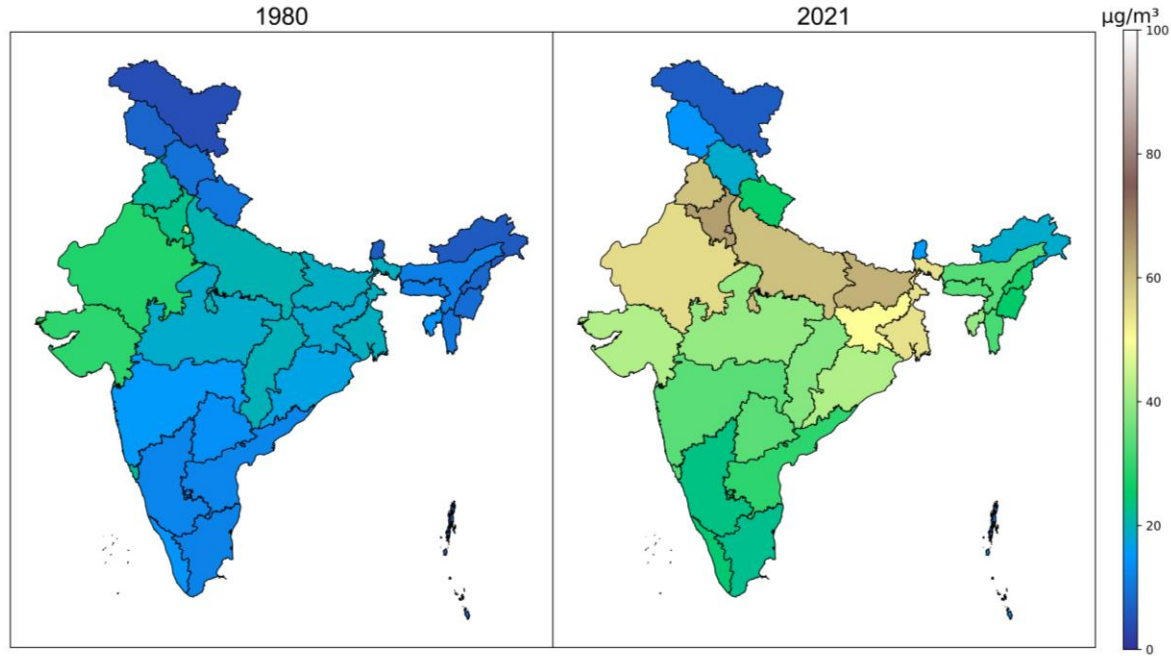
- Practice of designing and creating easy-to-communicate and understand graphic or visual representations of a large amount of complex data with the help of static, dynamic or interactive visual items.
- Representation of data using common graphics, such as charts, plots, infographics, and even animations.
- These visual displays of information communicate complex data relationships and data-driven insights in a way that is easy to understand.

Data Visualization

State/UTs	1980	2021
Andaman & Nicobar	7.486487	12.37268
Andhra Pradesh	12.36462	29.49085
Arunachal Pradesh	6.026727	18.54027
Assam	11.49391	33.5708
Bihar	19.13267	61.74638
Chandigarh	19.23599	48.21263
Chhattishgarh	19.75195	38.19519
Daman and Diu and Dadra and Nagar Haveli	23.5512	43.1923
Delhi	44.89746	84.48122
Goa	21.74392	33.1921
Gujarat	29.39484	42.43116
Haryana	22.72529	64.82074
Himachal Pradesh	9.43638	18.58373
Jammu and Kashmir	7.587756	14.50563
Jharkhand	18.18835	50.03474
Karnataka	12.36542	23.08367
Kerala	14.51138	24.32921
Ladakh	4.267032	6.379851

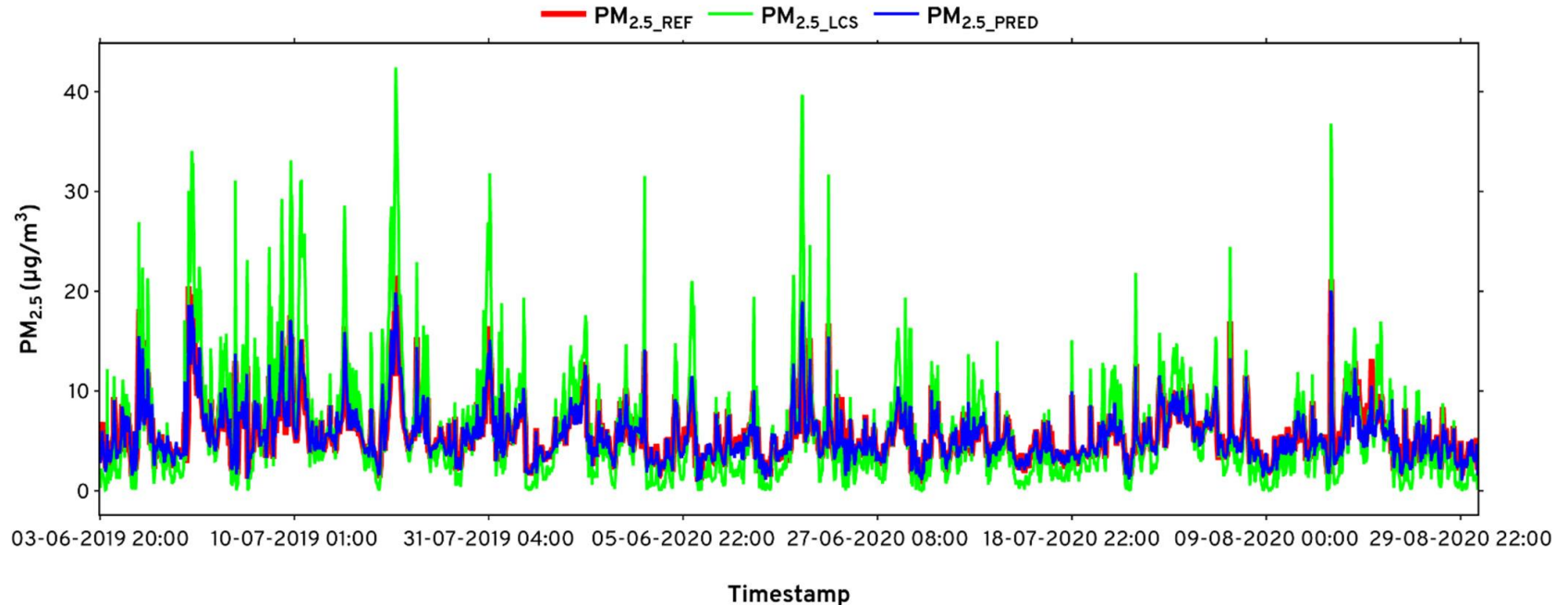
State/UTs	1980	2021
Lakshadweep	21.85739	28.80133
Madhya Pradesh	18.54075	39.5859
Maharashtra	15.57608	33.81927
Manipur	8.927957	24.98033
Meghalaya	11.60637	33.26539
Mizoram	10.3865	32.3418
Nagaland	8.103051	27.47781
Odisha	16.99674	42.2079
Puducherry	19.42465	31.91858
Punjab	21.37703	59.31624
Rajasthan	28.17039	55.10648
Sikkim	6.549081	15.5275
Tamilnadu	12.06234	22.36917
Telangana	14.04252	33.65671
Tripura	14.30304	39.93357
Uttar Pradesh	19.95941	59.729
Uttarakhand	10.46982	25.45957
West Bengal	19.29626	54.43993

Data Visualization



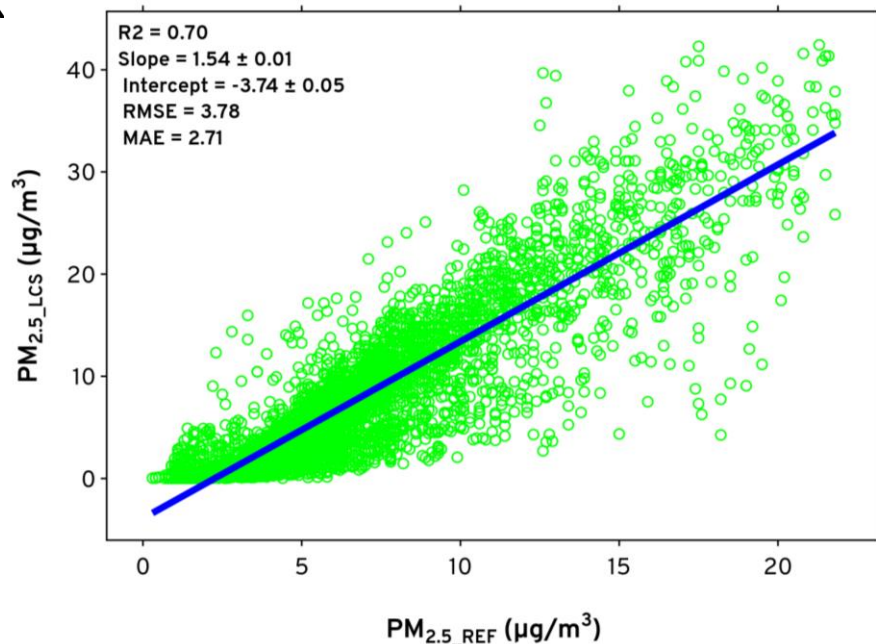
Line Chart

- Track the changes over the short and long period of time
- Compare the changes over the same period of time for different groups



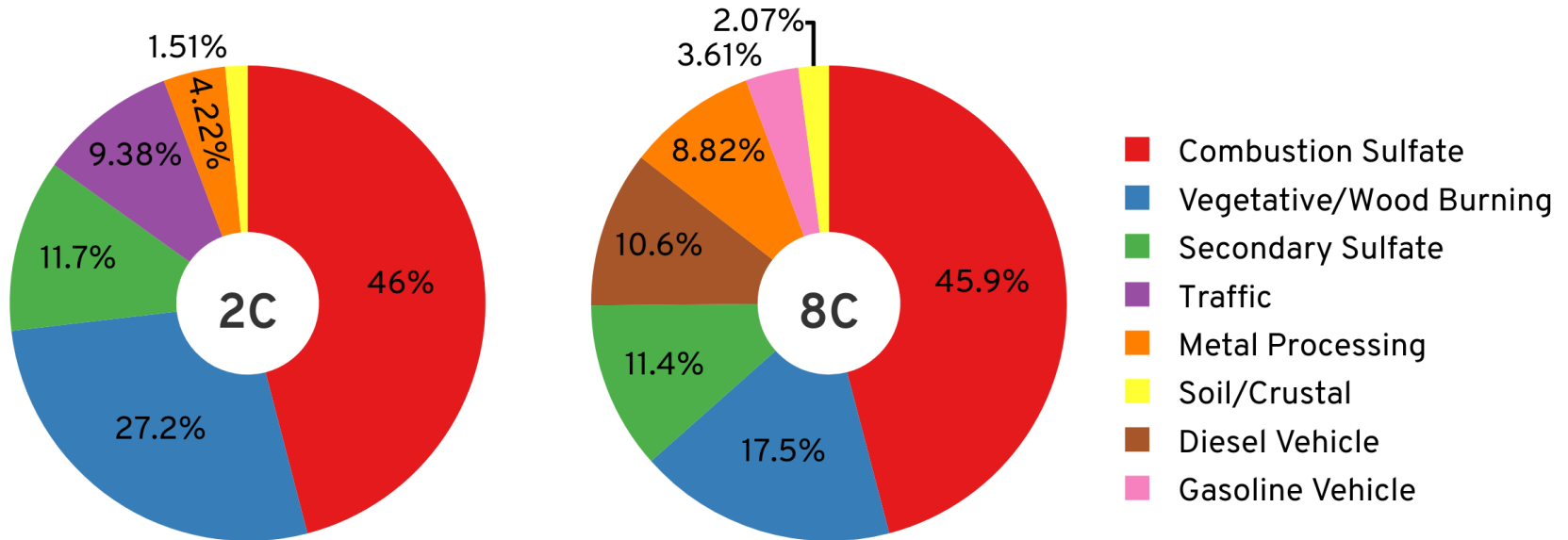
Scatter Plots

- Visually display the relationship and observe the nature between two variables
- Relationships observed can be positive or negative, non-linear or linear, strong or weak



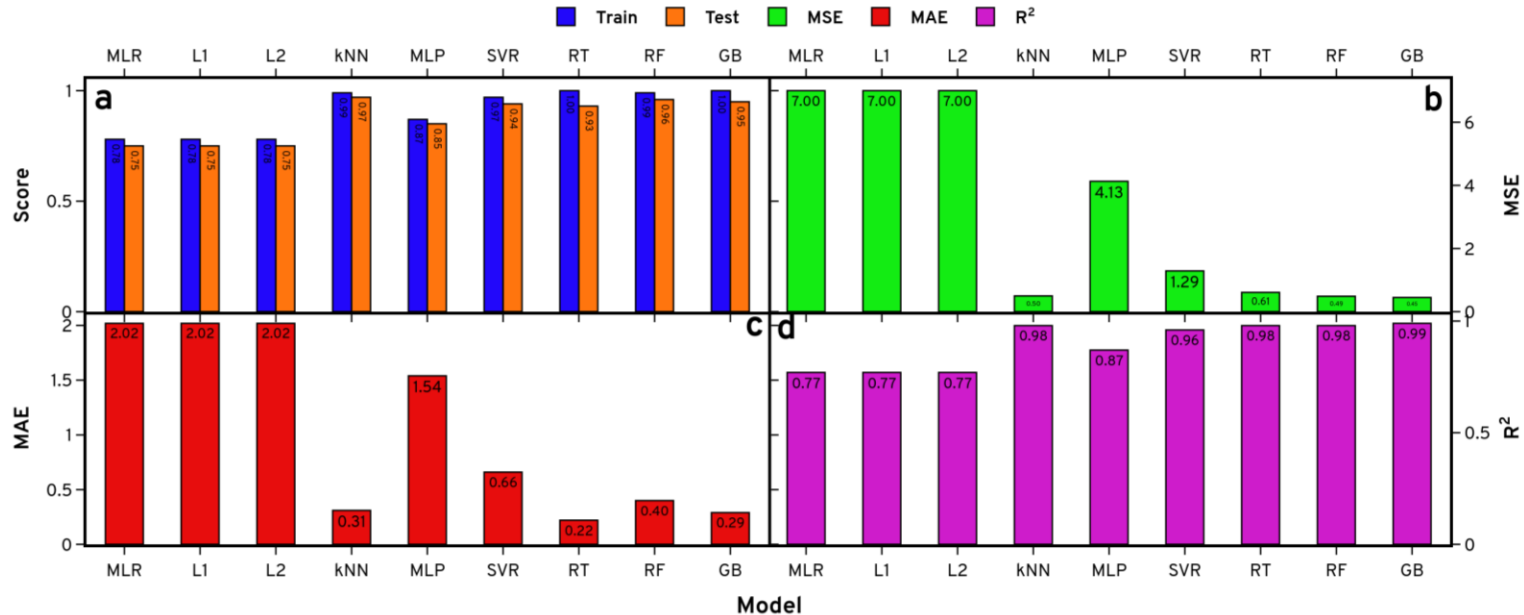
Pie Chart

- Used to represent the proportional data or relative data in a single chart
- Expresses a part-to-whole relationship in your data



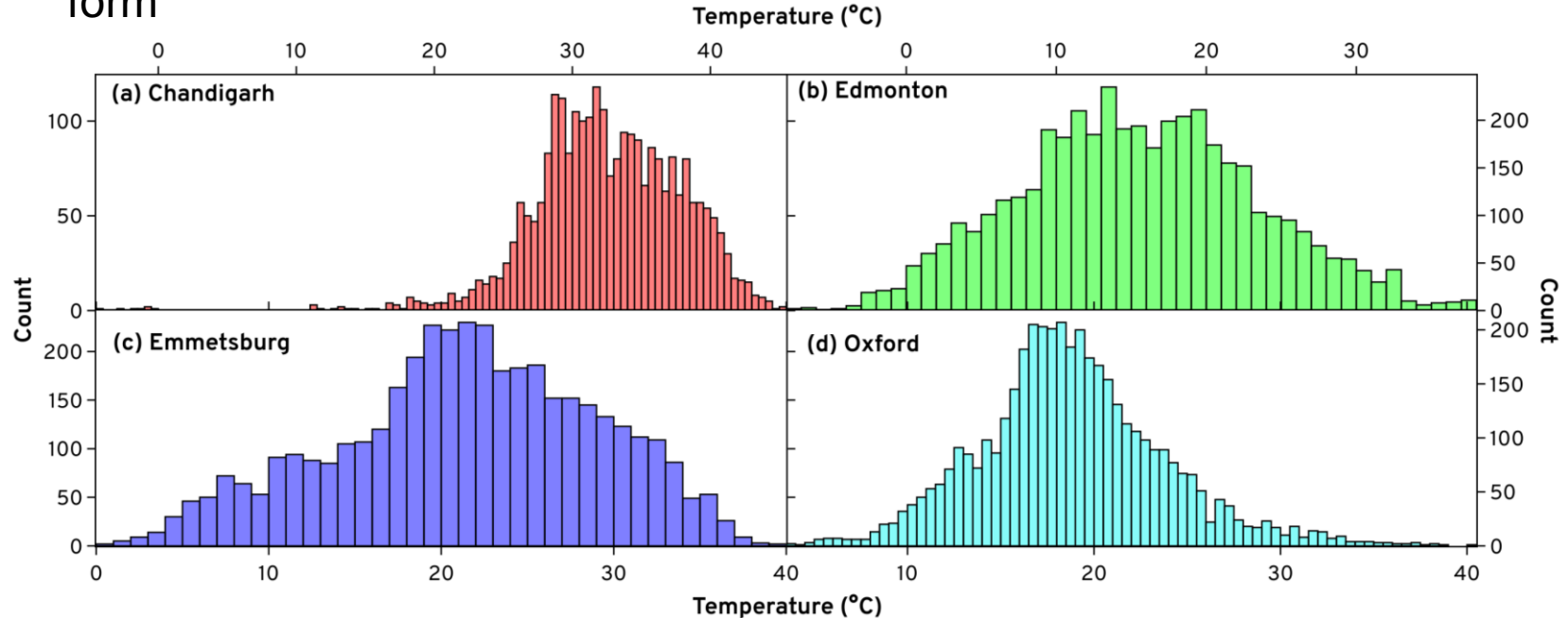
Bar Chart

- Comparing values between groups of data
- Comparisons between different variables are easy and convenient



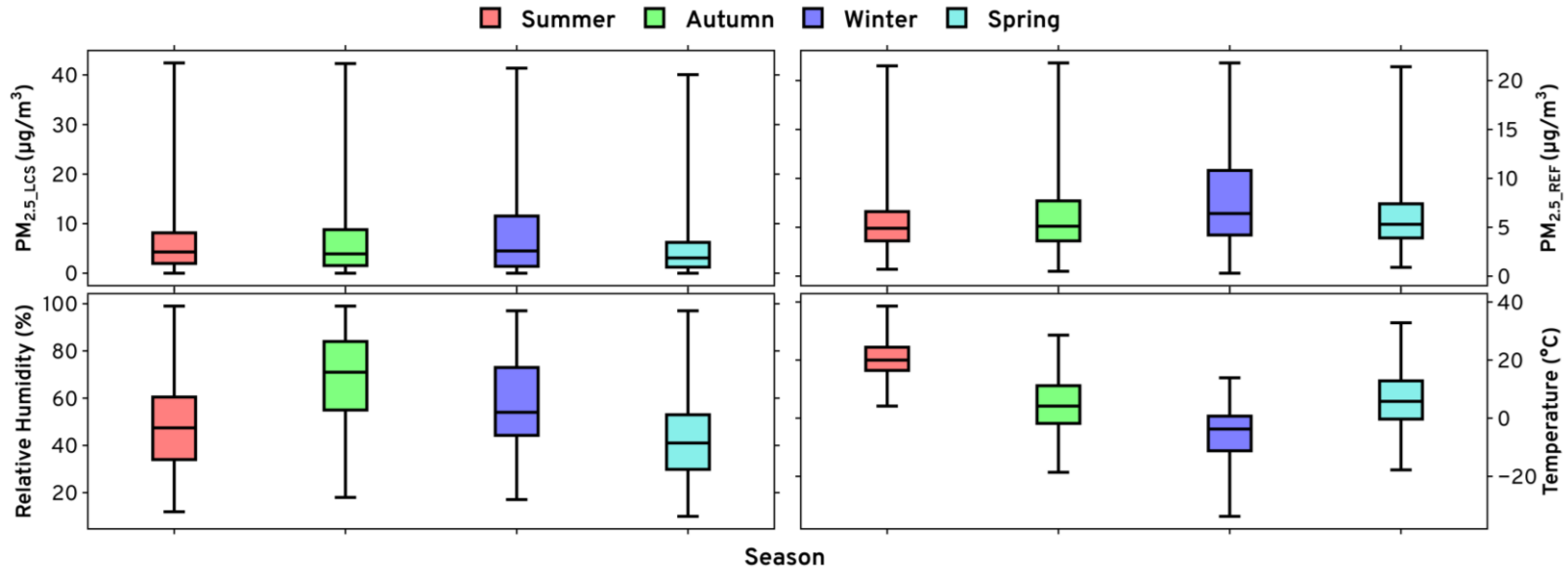
Histogram

- Summarize discrete or continuous data that are measured on an interval scale
- Illustrate the major features of the distribution of the data in a convenient form



Boxplot

- Show distributions of numeric data values to compare between multiple groups
- Provide general information about a group of data's symmetry, skew, variance, and outliers



Heatmap

- Graphical representation of data where values are depicted by color
- Depicts values for a main variable of interest across two axis variables as a grid of colored squares
- Each cell's color indicates the value of the main variable in the corresponding cell range

