Cheatsheet: Use of Generative AI for Data Analytics

Important Terms

Term	Description
Data Insights	Data insights are significant and insightful interpretations or findings that come from data analysis. They frequently entail identifying trends, patterns, correlations, or anomalies in a dataset.
Data Visualization	Data visualization is the use of graphical or visual representations to enhance understanding, analysis, and identification of patterns, trends, or insights within data.
Dashboard	Dashboard often consist of a collection of charts, graphs, tables, and other visual components used to present data in an easily understood visual format and are practical tools that help individuals and businesses make well-informed decisions.
Storytelling	Storytelling in data analytics involves transforming raw data into a compelling narrative that helps communicate insights, trends, and patterns in the data, making it easier for stakeholders to make informed decisions.
Hallucination	It refers to the generation of inaccurate or illogical information by generative AI models due to flawed training data, inappropriate model architectures, and inadequate evaluation methods.
AI Bias	AI bias refers to systematic inaccuracies or prejudices in AI systems' outcomes, affecting data collection, algorithm design, and model training. Addressing AI bias is crucial for fairness, equity, and ethical use of AI technologies.
Fairness	Fairness in AI involves treating all individuals or groups equitably without discrimination, avoiding favoritism based on race, gender, ethnicity, or socioeconomic status. This involves identifying biases, promoting transparency, and distributing AI benefits and risks equitably.

Generative AI Platforms/Tools used in this module

Task Performed	Generative AI Platform/Tool
Data Insights	hal9 ChatGPT
Data Visualization	Einblick Coulmn.ai Akkio
Dashboarding	ChartPixel
Storyteling	Akkio

Some Generic Prompts

Task	Prompt	Example
Get the statistical description of the dataset	Describe dataset	Describe dataset (upload the dataset on ChatCSV and then write a prompt)
Identify missing data	Write a <…> code to identify <…..> with missing values. Identify the attributes with missing data	Write a *python* code to Identify *the columns* with missing values (ChatGPT) Identify the attributes with missing data (ChatCSV)
Handling missing values	Write a <…> code to replace missing values with <…..> in the dataset Replace the missing values <…> in the <…> and save the updated dataset	Write a *python* code to replace missing values with *mean values* in the dataset. (ChatGPT) Replace the missing values with *the mean value* in the *Screen_size_cm column* and save the updated dataset. (ChatCSV)
Get different forms of visualizations from different types of attributes	Write a <…..>code to create <…..> plots for the attributes <…> against <…..>.	Write a Python code to perform the following actions: 1. Create *regression* plots for the attributes *CPU_frequency*, *Screen_Size_inch* and *Weight_pounds* against *Price*. 2. Create box plots for the attributes *Category*, *GPU*, *OS*, *CPU_core*, *RAM_GB* and *Storage_GB_SSD* against the attribute *Price*.

Author(s)

Dr. Pooja

Skills Network