

# Cheatsheet: Data Analytics and Generative AI

## Important Terms

Term	Description
<b>Generative AI</b>	Is a category of AI that focuses on creating new, synthetic data. Unlike traditional AI models that predict or classify, generative models generate entirely new data points, opening a realm of possibilities for data analytics.
<b>Data Augmentation</b>	Is a powerful technique for improving the performance of machine learning models, especially when the training data sets of the models are small or imbalanced.
<b>Data Preparation</b>	Is crucial in the data analytics journey. To prepare raw data for analysis, it must be cleaned, transformed, and arranged in a format that makes it easy for analytical tools to use.
<b>Database Querying</b>	Is the process of working with a database to extract relevant details for analysis. It includes interacting with a database and retrieving data that satisfies certain criteria using query languages, most often SQL (Structured Query Language).
<b>Data Q&amp;A</b>	Asking questions and getting answers about certain data sets or data analysis activities is called Q&A for data. With Q&A, you can accomplish data exploration, extract insights, and get a better comprehension of the underlying patterns and trends in the data.

## Generative AI Platforms/Tools Used in this Module

Task Performed	Generative AI Platform/Tools
Data Augmentation and Generation	<a href="#">DataRobot</a> <a href="#">Colab</a> <a href="#">ChatGPT</a> <a href="#">Bard</a> <a href="#">MOSTLY.AI</a> <a href="#">universaldata</a>
Data Preparation	<a href="#">ChatCSV</a> <a href="#">Tomat.AI</a>
Data Querying	<a href="#">SQLthroughAI</a> <a href="#">dbsensei</a>
Data Insights through QnA	<a href="#">Akkio</a>

## Some Generic Prompts

Task	Prompt	Example
Create dataset in a particular domain	Create <.....> dataset for <.....>	Create <i>patient</i> data set for the <i>symptoms of diabetes</i>
Create dataset with specific attributes and format	Create a dataset with attributes as <.....> in a <.....> format	Create a dataset with attributes as <i>temperature in Fahrenheit, Temperature Category, Humidity in percentage, Rain and Snow as categorical type in yes or no categories, Month, Year</i> in a <i>CSV</i> format
Get the insights through Q&A: Finding highest value within the data attribute	Identify the <.....> with the highest <.....>	Identify the products with the highest sales What are the top-selling products?
Get the insights through Q&A: See patterns of a data attribute over a period of time	How has <.....> changed over time?	How has the quantity ordered changed over time?
Identify missing data	Write a <...> code to Identify <.....> missing values. Identify the attributes with missing data	Write a <i>python</i> code to Identify <i>the columns</i> 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 <i>python</i> code to replace missing values with <i>mean values</i> in the dataset (ChatGPT) Replace the missing values with <i>the mean value</i> in the <i>Screen_size_cm</i> column and save the updated dataset (ChatCSV)
Join two tables	Write a SQL query to join <.....> with <.....> on the <.....> as a primary key	Write an SQL query to join <i>customer table</i> with the <i>product sales tables</i> on <i>product ID</i> as primary key (ChatGPT)

Task	Prompt	Example
	join <.....> with <.....> on <.....> as primary key	Join <i>customer table</i> with the <i>product sales table</i> on <i>product ID</i> as primary key (dbsensei)
Create database	Write a <.....> query to Create a database on <.....>, create a <.....> and insert values in these tables  Create a database on <.....>, create tables, insert values	Write a <i>SQL</i> query to Create a database on <i>sales</i> , create a <i>customer table</i> , a <i>sales table</i> , a <i>product table</i> and insert values in these tables (ChatGPT) Create a database on <i>customers and sales</i> , create tables, insert values (dbsensei)

## Author(s)

[Dr. Pooja](#)



# Skills Network