

# Cheat Sheet for comprehensive Google Data Analytics Professional Certificate

## Data Analysis Process

### 1. Ask

- **Define the Problem:** Clearly articulate the business task.
- **Identify Stakeholders:** List key stakeholders and their roles.
- **Questions to Ask:**
  - What is the problem you're trying to solve?
  - What are the goals of the project?
  - Who are the key stakeholders?

### 2. Prepare

- **Data Sources:**
  - Identify and gather relevant data from various sources (databases, APIs, spreadsheets).
- **Data Types:**
  - **Quantitative:** Numerical data (e.g., sales figures).
  - **Qualitative:** Descriptive data (e.g., customer feedback).
- **Data Organization:**
  - Use spreadsheets (Google Sheets, Excel) to organize and store data.
  - Ensure data is clean and free from errors.

### 3. Process

- **Data Cleaning:**
  - Remove duplicates.
  - Handle missing values (impute or remove).
  - Correct data entry errors.
- **Tools:**
  - **Google Sheets:** Use functions like `=UNIQUE()`, `=FILTER()`, `=SORT()`.
  - **SQL:** Use `SELECT`, `WHERE`, `GROUP BY`, `ORDER BY` to clean and process data.

- **R:** Use `dplyr` package for data manipulation (`filter()`, `select()`, `mutate()`).

#### 4. Analyze

##### - **Descriptive Statistics:**

- Mean, median, mode.
- Range, variance, standard deviation.

##### - **Data Visualization:**

- **Google Sheets:** Use charts (bar, line, pie) and pivot tables.
- **R:** Use `ggplot2` for advanced visualizations.
- **Tableau:** Create interactive dashboards.

##### - **Statistical Tests:**

- **T-tests:** Compare means of two groups.
- **Chi-square tests:** Test for independence between categorical variables.

#### 5. Share

##### - **Visual Reports:**

- Create clear and concise visual reports.
- Use storytelling techniques to present findings.

##### - **Tools:**

- **Google Slides:** Create presentations.
- **RMarkdown:** Generate HTML, PDF reports.
- **Tableau Public:** Share interactive dashboards.

#### 6. Act

##### - **Recommendations:**

- Provide actionable insights based on analysis.
- Suggest next steps and potential improvements.

##### - **Feedback Loop:**

- Continuously gather feedback and iterate on the analysis.

## Tools and Techniques

### Google Sheets

#### - **Basic Functions:**

- `=SUM()`, `=AVERAGE()`, `=COUNT()`.
- `=IF()`, `=VLOOKUP()`, `=INDEX()`.

#### - **Advanced Functions:**

- `=ARRAYFORMULA()`: Apply a formula to an entire column.
- `=QUERY()`: Run SQL-like queries on data.

#### - **Shortcuts:**

- `Ctrl + C`, `Ctrl + V`: Copy and paste.
- `Ctrl + Z`: Undo.
- `Ctrl + Shift + Enter`: Array formula.

### SQL

#### - **Basic Queries:**

- `SELECT column_name FROM table_name;`
- `SELECT * FROM table_name WHERE condition;`

#### - **Advanced Queries:**

- `SELECT column_name, COUNT(*) FROM table_name GROUP BY column_name;`
- `SELECT column_name FROM table_name ORDER BY column_name DESC;`

#### - **Joins:**

- `INNER JOIN`, `LEFT JOIN`, `RIGHT JOIN`, `FULL OUTER JOIN`.

### R

#### - **Basic Syntax:**

- `data <- read.csv("file.csv")`
- `summary(data)`

#### - **Data Manipulation:**

- `library(dplyr)`
- `data %>% filter(condition) %>% select(columns)`

#### - **Visualization:**

- `library(ggplot2)`

- `ggplot(data, aes(x=column1, y=column2)) + geom_point()`

### Tableau

#### - **Connecting to Data:**

- Connect to various data sources (Excel, SQL, Google Sheets).

#### - **Creating Visualizations:**

- Drag and drop dimensions and measures.
- Use filters and calculated fields.

#### - **Publishing:**

- Publish dashboards to Tableau Public or Tableau Server.

### Tips and Tricks

#### Data Cleaning

- **Identify Missing Data:** Use `=COUNTBLANK()` in Google Sheets.
- **Remove Duplicates:** Use `=UNIQUE()` or `Remove Duplicates` in Excel.
- **Normalize Data:** Ensure consistent formatting (e.g., date formats).

#### Data Visualization

#### - **Choose the Right Chart:**

- **Bar Chart:** Compare categories.
- **Line Chart:** Show trends over time.
- **Pie Chart:** Display parts of a whole.
- **Color Palettes:** Use consistent and accessible color schemes.
- **Annotations:** Add labels and annotations to highlight key points.

#### Reporting

- **Storytelling:** Use narratives to guide stakeholders through findings.
- **Clarity:** Keep reports concise and focused.
- **Interactive Elements:** Use filters and drill-downs in Tableau for interactivity.

## Examples

### Google Sheets Example

```
=SUM(A1:A10) // Sum of values in cells A1 to A10  
=VLOOKUP("Apple", A1:B10, 2, FALSE) // Find "Apple" in column A and  
return corresponding value in column B
```

### SQL Example

```
SELECT CustomerID, SUM(OrderAmount)  
FROM Orders  
GROUP BY CustomerID  
ORDER BY SUM(OrderAmount) DESC;
```

### R Example

```
library(dplyr)  
data %>%  
  filter(Year == 2023) %>%  
  select(Product, Sales) %>%  
  arrange(desc(Sales))
```

## Conclusion

- **Continuous Learning:** Stay updated with the latest tools and techniques.
- **Practice:** Regularly practice with real-world datasets.
- **Collaboration:** Work with others to improve analysis and reporting skills.

By Ahmed Baheeg Khorshid

ver 1.0