**5 high-value use cases with prompt templates** tailored for data analytics:

### 1. Data Cleaning & Preprocessing

* 🤖 Prompt template:  
  "Clean this dataset by handling missing values, removing duplicates, and normalizing column names. Output Python (pandas) code only."
* 🔧 Use case: Automating repetitive data wrangling steps.

### 2. Exploratory Data Analysis (EDA)

* 📊 Prompt template:  
  "Perform an EDA on this dataset: summarize key statistics, detect outliers, and suggest visualizations. Provide Python code using pandas & matplotlib."
* 🔧 Use case: Quick overview of dataset health & trends.

### 3. Feature Engineering

* 🧩 Prompt template:  
  "Suggest 10 potential new features from this dataset for a predictive model. Explain why they might be useful."
* 🔧 Use case: Enhancing ML model performance with better features.

### 4. SQL Query Generation

* 💾 Prompt template:  
  "Write an SQL query to [task, e.g., get the top 10 customers by revenue last quarter]. Optimize for readability and efficiency."
* 🔧 Use case: Fast query generation & optimization.

### 5. Data Storytelling & Insights

* 📈 Prompt template:  
  "Explain the main insights from this dataset in 5 bullet points for a business audience. Keep it concise and non-technical."
* 🔧 Use case: Turning raw data into actionable business insights.

## 🧰 Master Data Analytics Prompt

\*"You are a data analyst assistant. I will provide you with a dataset (or description of it). Your job is to:

1. [CLEAN] → Handle missing values, duplicates, formatting issues.
2. [EDA] → Summarize dataset (stats, distributions, outliers). Suggest 3–5 visualizations.
3. [FEATURES] → Recommend useful new features for modeling.
4. [SQL] → Write SQL queries for key business questions.
5. [INSIGHTS] → Summarize findings in 5 bullet points for a business audience.

Output must include:

* Python code (using pandas, matplotlib/seaborn).
* SQL queries.
* Plain-language insights.  
  Keep explanations concise and structured."\*

## ⚡ Example 1 (Sales Data)

"Dataset: Online retail transactions (columns: customer\_id, product\_id, purchase\_date, revenue, region)."

👉 Output expectations:

* 🧹 Python code cleaning missing purchase\_date, removing duplicate transactions.
* 📊 EDA summary: avg. revenue per region, top-selling products, purchase frequency.
* 🧩 Features: customer lifetime value, recency/frequency, product categories.
* 💾 SQL: top 10 customers by revenue, monthly revenue trend.
* 📈 Insights: "Region X drives 45% of revenue", "High-value customers buy monthly", etc.

## ⚡ Example 2 (HR Data)

"Dataset: Employee info (columns: employee\_id, department, salary, tenure, performance\_score)."

👉 Output expectations:

* 🧹 Python code: fix salary outliers, normalize department names.
* 📊 EDA summary: distribution of tenure, salary vs performance.
* 🧩 Features: tenure buckets, salary bands, performance trends.
* 💾 SQL: top departments by avg. performance, attrition risk.
* 📈 Insights: "Sales has high salaries but lower performance variance."

**reusable one-liner master prompt** you can copy-paste every time:

⚡ **One-Liner Master Prompt**  
"Act as a data analyst assistant. Given this dataset: [insert dataset or description], perform the following: (1) Clean and preprocess data, (2) Summarize EDA with key stats and 3–5 suggested visualizations, (3) Propose useful new features, (4) Write 2–3 SQL queries for business questions, (5) Provide 5 plain-language insights for non-technical stakeholders. Output structured Python code (pandas/matplotlib), SQL queries, and concise business insights."

**Master Prompt + insights output template** (so you just drop in data and it auto-fits into a clean structure)?

⚡ **Final Master Prompt**  
\*"Act as a data analyst assistant. Given this dataset: [insert dataset or description], perform the following and present results in this exact structure:

1. **Data Cleaning & Preprocessing** → Python (pandas) code.
2. **Exploratory Data Analysis (EDA)** → Key stats, anomalies, 3–5 suggested visualizations + Python code.
3. **Feature Engineering** → 3–5 new features with justifications + Python code.
4. **SQL Queries** → 2–3 useful business queries.
5. **Business Insights** → 5 concise plain-language insights for non-technical stakeholders.

Output must follow this template:

* 🧹 Data Cleaning → code block
* 📊 EDA → bullet points + code block
* 🧩 Features → bullet points + code block
* 💾 SQL → SQL code block
* 📈 Insights → 5 bullets."\*