

## Mermaid code:

flowchart TD

%% Main Branch (rectangles)

A[Raw Data Sources] --> B[Data Ingestion]

B --> C[Data Preparation]

C --> D[Modeling]

D --> E[Evaluation]

E --> F[Deployment]

F --> G[Monitoring & Maintenance]

%% Sub-branches (flattened ellipses)

B --> B1([Databases])

B --> B2([APIs])

B --> B3([Files])

C --> C1([Cleaning])

C1 --> C1a([Missing Values])

C1 --> C1b([Duplicates])

C --> C2([Feature Engineering])

C2 --> C2a([Feature Selection])

C2 --> C2b([Encoding])

C --> C3([Train/Test Split])

D --> D1([Choose Algorithm])

D --> D2([Train Model])

D --> D3([Hyperparameter Tuning])

E --> E1([Metrics])

E1 --> E1a([Accuracy])

E1 --> E1b([Precision/Recall])

E1 --> E1c([RMSE])

F --> F1([API/Service])

F --> F2([Web App])

F --> F3([Batch Jobs])

G --> G1([Performance Tracking])

G --> G2([Retraining])

%% Styling

class A,B,C,D,E,F,G main;

classDef main fill:#87CEFA,stroke:#000,stroke-width:2px,font-size:16px;

## Code explanation:

### 1. Define the pipeline flow:

Flowchart TD tells Mermaid that we want a top-down flowchart (TD = top to bottom).

### 2. Main pipeline steps (rectangles):

- Each main step is a rectangle [ ].
- Arrows --> show the direction of flow (step order).

So the main pipeline backbone is:

Raw Data → Data Ingestion → Data Preparation → Modeling → Evaluation → Deployment → Monitoring

### 3. Sub-branches (flattened ellipses):

- Sub-steps are connected to their main step.
- They use ([ ]), which creates flattened ellipses, different from rectangles.

Example: Data Ingestion (B) has sub-branches for Databases, APIs, and Files.

This pattern is repeated for other main steps like Data Preparation, Modeling, Evaluation, etc.

### 4. Styling the main branches:

- class A,B,C,D,E,F,G main; → assigns the class main to all main rectangles.
- classDef main ... → defines how main class looks:
  - a. fill:#87CEFA → light blue background
  - b. stroke:#000 → black border
  - c. stroke-width:2px → thicker border for emphasis
  - d. font-size:16px → slightly larger font for readability

Finally, key points:

- Rectangles → main pipeline steps (backbone).
- Flattened ellipses → sub-steps/details under each main step.
- Arrows → show the sequence of steps.
- Class styling → makes main steps visually distinct.