

# Week 1: Planning Phase

## SWOT Analysis

### Strengths

- Skills in Python and Public Tableau for data analysis
- Ability to clean and analyze structured datasets
- Knowledge of data visualization

### Weaknesses

- No exposure to advanced statistical models or ML
- Need for more practice in writing efficient SQL queries

### Opportunities

- Data-driven decision-making is in high demand across industries
- Public datasets available for real-world (government data)

### Threats

- Incomplete or inconsistent data
- Interpreting results incorrectly due to lack of domain knowledge

## Project Proposal

### Project Objective

- *This project aims to analyze unemployment annual trends for counties and cities using Python and Public Tableau.*

### Scope & Timeline

- Define data sources:
  - CSV files
  - SQL databases
  - JSON from web scraping
- *State of California Unemployment data will be imported from a CSV file, cleaned and analyzed with Python and Public Tableau visualizations.*

### Timeline

- Week 1: Planning and defining project scope

- Week 2: Data collection and preprocessing
- Week 3: Data analysis using Python
- Week 4: Creating visual reports in Public Tableau
- Week 5: Final improvements and project documentation

### Expected Outcome

- *A Public Tableau dashboard displaying key Unemployment trends, by Counties and Cities over time.*

### Risks & Mitigation Strategies

- *If the dataset contains missing values, they will be handled using interpolation or removal techniques. If data is inconsistent, data cleaning will be performed using Python's pandas library.*

## Week 2: Analysis Phase

### Prompt

In this phase, you will define your project's technical needs, choose a development approach, and create system models to understand data flow and operations.

## 1. System Requirements Document

- **Input:**
  - <https://catalog.data.gov/dataset/local-area-unemployment-statistics-laus-annual-average>
  - CSV
  - Ensure each county and city has correct name and numeric value
- **Processing:**
  - Data cleaned using Python Pandas library to search for missing values, incorrect names, and dates
  - Data transformed with Python Pandas to ensure all data inputs ready for Public Tableau
  - Data analyzed within Public Tableau and via dashboard
- **Output:**
  - Public Tableau visualizing Exploratory Data Analysis and Dashboard for Descriptive Analysis on county and city Unemployment Trends in the golden state of California from 1990 to 2025.

## 2. Development Methodology Justification\*\* (*Short report*)

- Identify the chosen methodology (Agile, CRISP-DM, etc.) and justify its relevance.
- Outline key milestones and roles (if applicable).
  - *Example:* A brief explanation of why Agile was chosen for iterative dashboard development.

### 3. UML Diagrams\*\* (*PDF or PNG format*)

- **Use Case Diagram:** Show user interactions with the system.
- **Class Diagram:** Define key data objects and their attributes.
  - *Example:* A Use Case Diagram illustrating how analysts retrieve and filter sales data.

### 4. Data Flow Diagrams (DFDs)\*\* (*PDF or PNG format*)

- **DFD Level 0:** High-level overview of data movement.
- **DFD Level 1:** Detailed breakdown of data flow between components.
  - *Example:* A DFD showing how transaction data moves from a SQL database to a Power BI dashboard.

### 5. Security and Storage Plan\*\* (*Short report*)

- Describe how data will be stored (local, cloud, or hybrid).
- Identify security risks and planned safeguards (e.g., encryption, API security).
  - *Example:* A report explaining how sensitive customer data will be encrypted and backed up.

In [ ]:

## Week 1: Planning Phase

## Week 1: Planning Phase

## Week 1: Planning Phase