Data Analytics Graduation Project Proposal

Project Title: California Wildfire Damage Analysis

Client: Digital Egypt Pioneers Initiative (DEPI)
Team Members: Lamia Eslam Eldeeb, Reem Mostafa Mohamed Kamal Mohamed, Mikhail Samuel Ibrahim Samuel, Mayan Essam Ibrahim Gad, Sohila Ragab Abdelazeem
Due Date: April 2025

Executive Summary

This project analyzes **California Wildfire Damage Dataset (2014–February 2025)** to uncover trends, financial losses, and wildfire impacts across locations. Using data analytics and visualization, we aim to support disaster management, risk assessment, and policy-making. The final deliverables include a cleaned dataset, business insights, forecasting trends, and a Tableau dashboard for visualization.

Proposed Solution & Value Addition

1. **Comprehensive Analysis:** Identify key trends, geographical impacts, and wildfire causes.
2. **Data Preprocessing & Cleaning:** Clean and enhance the dataset using Python (Pandas, NumPy, Seaborn, Matplotlib) and add new valuable columns.
3. **Data-Driven Insights:** Store and query the dataset using SQL Server for analysis.
4. **Visualization for Decision-Making:** Develop an interactive Tableau dashboard for insights.

Project Timeline & Deliverables

Week 1: Build Data Model, Data Cleaning, and Preprocessing using Python (Pandas, NumPy, Seaborn, Matplotlib), SQL Server

Deliverables:

- Cleaned dataset ready for analysis.
- Data preprocessing notebook documenting transformations.

Week 2: Analysis Questions Phase using SQL, Python (Pandas, Matplotlib)

Deliverables:

- Structured set of analysis questions to be answered using the dataset.

Week 3: Forecasting Questions Phase using Python (Scikit-learn, Pandas, Matplotlib)

Deliverables:

- Visualization plots answering forecasting questions.

Week 4: Visualization Dashboard & Final Presentation using Tableau, Microsoft Power point

Deliverables:

- Tableau visualization dashboard for insights.
- Final report and presentation.

Technical Specifications

- **Tools & Technologies:** Python (Pandas, NumPy, Seaborn, Matplotlib, Scikit-learn), SQL Server, Tableau.
- **Data Model:** ERD with one-to-one relationships across three views (Original, Business, Human).
- **Key Metrics:**
    - Wildfire incident trends
    - Financial losses over time
    - Geographical impact (counties most affected)
    - Cause analysis (lightning, human activity, unknown)
    - Fatalities and injuries by location & season

Financial Considerations

This is an academic research project with no external funding. However, the insights can benefit government agencies, environmental organizations, and disaster management teams in policy-making and resource allocation.

Conclusion

This project will provide data-driven insights into California wildfires, helping researchers, policymakers, and emergency management teams understand trends and financial impacts. The Tableau dashboard and final report will serve as valuable resources for wildfire mitigation and policy decisions.