Revised 3-Day PhD-Level Project Plan

Title: "AI & Nuclear Intelligence: A Multimodal Study of Energy Production, Safety, and Policy Readiness"

Day 1: Data Discovery + Warehouse Foundation + SQL Exploration

GOAL: Understand dataset, create a mini Data Warehouse schema, extract insights using SQL

Step 1: Understand & Document Datasets

- Define dimensions and facts from each CSV
- Key files: nuclear_energy_overview_eia.csv, uranium_purchase_price_us.csv, power_plant_database_global.csv, rates_death_from_energy_production_per_twh.csv
- Learn: Star vs Snowflake Schema (Kimballs Toolkit, YouTube DW Basics)

Step 2: Design Your Data Warehouse Schema

- Fact_EnergyMetrics nuclear_output, death_rate, uranium_price
- Dimensions: Country, Energy Type, Time, Company, Reactor
- Use ERD tool or draw schema manually

Step 3: Ingest & Query Using SQL

- Load data using Python or SSMS
- Sample Questions:
  1. Total nuclear energy by continent over time
  2. Uranium prices by year
  3. Top 10 companies by reactors
  4. Highest death rate per TWh
- Learn: SQL Joins, GROUP BY, Window Functions

Day 2: Python Analysis + Machine Learning (Beginner-Friendly)

GOAL: Explore with pandas, apply KMeans Clustering and simple Regression

Step 1: Clean + Merge Data Using Python

- Tasks: clean nulls, merge by country/year, generate trends
- Questions:
  1. Top nuclear energy producers
  2. Uranium price trends
  3. Outlier detection in safety/production
- Learn: pandas, matplotlib, seaborn (Kaggle Python course)

Step 2: Build Machine Learning Models

- KMeans Clustering on: output, price, death rate, reactors
- Optional: Linear Regression to predict uranium price
- Learn: KMeans, Regression, train/test split (StatQuest, Google ML Crash Course)

Day 3: Power BI + Research Summary Dashboard

GOAL: Visual storytelling + academic-style summary

Step 1: Load Data into Power BI

- Visual Ideas:
  - Line chart: Nuclear output vs time
  - Map: Reactors/plants by country
  - Cluster result with colored groups
  - KPI Cards for risk factors
- Learn: Power Query, visuals, slicers (Power BI Udemy Course)

Step 2: Compile Research Summary (Jupyter/PDF)

- Title Page, Abstract, Methods (SQL, Python, DW, ML), Results (Power BI), Recommendations

Folder Structure

ai_nuclear_project/

data/raw & cleaned

sql/

python/

powerbi/

docs/

README.md