Project Planning

Presented by -
Sanya Syed

Project Plan Overview:

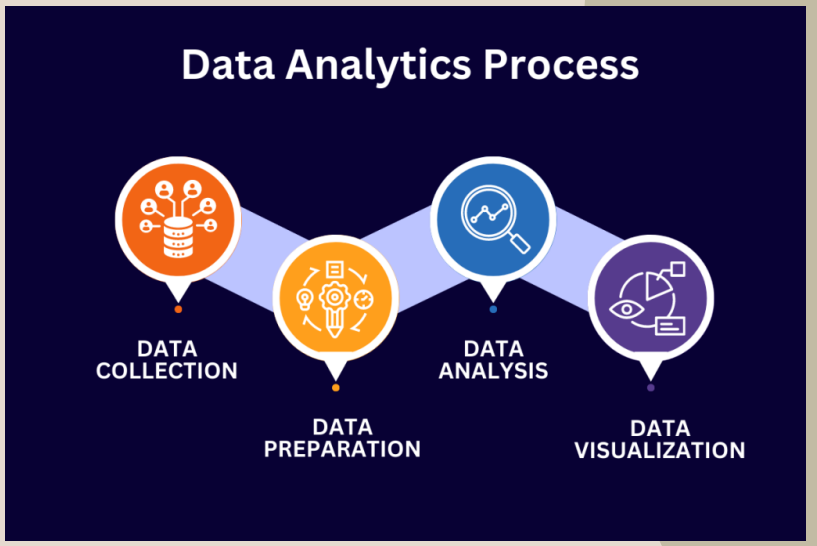
- Aim: What You Want to Do?

  - **Simple Project**: Basic data analysis to find insights.

  - **Intermediate Project**: In-depth data analysis for actionable insights.

Project Plan Overview:

- Objective: Analyze a dataset to find basic insights.
- Stage 1: Data collection or access existing dataset Tools: Excel, Google Sheets, CSV files, Kaggle Datasets
- Stage 2: Data cleaning, preparation, and integration Tools: Python (Pandas), SQL (Postgres), Talend, Power Query (Power BI)
- Stage 3: Basic visualizations and summary statistics Tools: Excel (charts), Python (Matplotlib, Seaborn)



Advance Project:

- Objective: Perform a deeper analysis and derive actionable insights.
- Stage 1: Define the business problem and analytical goals Tools: Google Docs, Microsoft Word, Jira (for task planning)
- Stage 2: Data acquisition (gathering multiple data sources if needed) Tools: SQL (Postgres, MySQL), Python (for API calls, web scraping)
- Stage 3: Data cleaning, preparation, and integration Tools: Python (Pandas), SQL (Postgres), Talend, Power Query (Power BI)
- Stage 4: Feature engineering and transformation Tools: Python (Scikit-learn, Pandas), SQL
- Cont.…

Advance Project:

- Stage 5: Exploratory data analysis (EDA) with advanced techniques Tools: Python (Matplotlib, Seaborn), Jupyter Notebook
- Stage 6: Statistical analysis or model building (if relevant) Tools: Python (Scikit-learn, Statsmodels), R, Excel (for simpler models)
- Stage 7: Advanced visualizations and storytelling with data Tools: Power BI, Tableau, Python (Plotly, Seaborn)
- Stage 8: Draw conclusions, derive insights, and provide recommendations Tools: PowerPoint, Google Slides, Word for reporting
- Stage 9: Final report and presentation Tools: Power BI (dashboards), Tableau, PowerPoint

Replication as a Learning Method:

- Key to reinforcing knowledge.
- Helps to understand processes by repeating and refining tasks.
- Use for improving problem-solving skills in projects.

Resources for Learning & Development:

- Online Platforms:
  - Kaggle: For datasets, competitions, and community discussions.
  - Git: For version control and collaboration on projects.
- Additional Learning Resources:
  - Zoomcamp: Free online bootcamp.
  - Unpaid/Paid Courses: For structured, in-depth learning

Potential Blockers in Data Projects:

- Common Challenges:
  - Environment Issues: Unable to open files.
  - Error Messages: Needing to read and search for fixes.
  - Version Conflicts: Mismatches in Python or package versions.
  - Expertise Levels: Varying levels of understanding can cause delays.
- Solution:
  - Use official documentation (e.g., Python docs, Postgres docs) to resolve syntax or functionality questions.

Key Takeaways for Success:

- **Effort:** Consistent work leads to progress.
- **Use Cases:** Aim for real-world applicability, particularly in corporate environments.
- **Focus on Outcomes:** Provide insights and actionable recommendations through your analysis.

Q & A