# Minor Project Guidelines: Exploratory Data Analysis (EDA) on Diwali Sales Dataset

## 📌 Project Title
Exploratory Data Analysis on Diwali Sales Dataset

## 🎯 Objective
The goal of this minor project is to perform an in-depth Exploratory Data Analysis (EDA) on a Diwali Sales dataset to gain insights into consumer behavior, product performance, and regional sales trends.

## 🛠️ Tools & Libraries
- Python (Jupyter Notebook or Google Colab)
- Pandas
- NumPy
- Matplotlib
- Seaborn

## 📋 Step-by-Step Instructions
1. Download the Dataset
• From Kaggle: Diwali Sales Data
• Link: https://www.kaggle.com/datasets

2. Data Import & Export
• Import the dataset using pandas
• After cleaning, export the final version as 'diwali_cleaned.csv'

3. Data Cleaning
• Remove unnecessary columns

• Rename columns for clarity
• Handle missing/null values
• Remove duplicate rows
4. Data Preprocessing
• Convert data types where needed (e.g., Amount → numeric)
• Remove or handle outliers
• Encode categorical variables if necessary

5. EDA (Exploratory Data Analysis)
Create at least 10 different visualizations, including:
- Gender-wise purchase
- Age group vs purchase amount
- State-wise sales (bar plot)
- Occupation vs purchase
- Product category popularity
- Marital status vs purchase
- Top 10 cities by sales
- Purchase amount distribution (histogram)
- Boxplot: Purchase by gender - Any other meaningful chart

## Deliverables

1. Cleaned CSV file: diwali_cleaned.csv
2. Final .ipynb file (Jupyter Notebook) OR Google Colab link 3. One-page PDF or DOCX Project Report (Mandatory), including:
- Project title & objective
- Tools and libraries used
- Screenshots of key visualizations
- Summary of insights - Challenges faced and solutions

## Assigned By

Mr. Ayush Shrivastav