

- 
-  **Exam Type:** Theory + Practical  
 **Duration:** 6 Hours  
 **Dataset:** A gigantic CSV dataset: `loan_applications.csv` (Generate from an AI tool)
- 

### Project Problem Statement:

You are provided with a dataset (`loan_applications.csv`) containing ~5000 records with the following fields:

- Customer\_ID
- Age
- Income
- Loan\_Amount
- Credit\_Score
- Loan\_Term (months)
- Default\_Status (Yes/No)

Your goal is to **explain key statistical concepts in theory** and then apply **Python-based statistical & probability analysis** to evaluate loan default risk.

---

### Exam Tasks

#### Part A – Theory (Short Questions)

1. Explain **Mean, Median, Mode** in the context of customer income.
2. Differentiate between **Standard Deviation** and **Variance** using loan amounts.
3. What is a **Random Variable**? Give one example from the dataset.
4. Explain **Conditional Probability** in terms of loan defaults.
5. Define **Bayes Theorem** and mention how banks can apply it.
6. Differentiate between **Empirical Probability** and **Theoretical Probability** with examples.
7. What is a **Poisson Distribution**? Give a business example.

8. Write a short note on **Eigenvalues and Eigenvectors** in data analysis.
- 

### **Part B – Practical (Python Programming)**

#### **Step 1: Central Tendency & Dispersion**

- Find **mean, median, and mode** of Income.
- Calculate **range, variance, and standard deviation** of Loan\_Amount.

#### **Step 2: Probability & Events**

- Compute probability of loan default.
- Create a **contingency table** between Default\_Status and Credit\_Score (categorized into ranges).
- Compute **conditional probability**:  $P(\text{Default} | \text{Credit\_Score} < 600)$ .

#### **Step 3: Distributions & Visualization**

- Plot a **Histogram** of Credit\_Score with a Gaussian curve.
- Check **Skewness and Kurtosis** of Loan\_Amount.
- Draw a **Q-Q Plot** for Income.

#### **Step 4: Linear Algebra Application**

Take the first 5 customers' [Income, Loan\_Amount] as **vectors**.

- Perform **dot product** between two customer vectors.
  - Find **Norm 2** of a customer's financial vector.
  - Calculate the **angle** between two customers' vectors.
- 

### **Deliverables (to be submitted)**

1. **Python file / Jupyter notebook** with all calculations.
  2. **PDF/Word report** with:
    - Short theory answers (Part A).
    - Screenshots/outputs of Part B practical tasks.
    - 3–5 insights (e.g., "Customers with Credit Score < 600 have 65% chance of default").
- 

### **Practical Exam** Mathematics & Advanced Statistics

---

**BRING ON YOUR CODING ATTITUDE**