

# Guidelines for Project Distribution & Evaluation in Statistics

## CA Distribution

- **CA 1 (50 Marks)** → *Descriptive Statistics & EDA*
  - Students select a dataset and perform **data cleaning, descriptive statistics, and visualization**.
  - Focus: foundations and exploratory insights.
- **CA 2 (50 Marks)** → *Statistical Modelling & Inference*
  - Students apply **hypothesis testing, regression, probability models, or correlation analysis**.
  - Focus: applying inferential statistics and interpreting results.
- **CA 3 (50 Marks)** → *Final Report & Presentation*
  - Students consolidate work into a **comprehensive report + presentation**.
  - Should include: **dataset intro, methodology, results, insights, and recommendations**.
  - Focus: integration of all methods, communication of results.

---

## Rubrics for Evaluation (50 Marks Each)

### CA 1: Descriptive Statistics & EDA (50 Marks)

Criterion	Marks	Description
Dataset Selection & Cleaning	10	Relevance, handling missing/outliers
Descriptive Statistics	15	Accuracy of mean, median, mode, variance, skewness, kurtosis
Data Visualization	10	Quality and relevance of graphs (histogram, scatter, boxplot, etc.)
Interpretation of Results	10	Meaningful insights, not just numbers

Criterion	Marks	Description
Report Presentation	5	Neatness, structure, timely submission

---

### CA 2: Statistical Modelling & Inference (50 Marks)

Criterion	Marks	Description
Method Selection	10	Appropriate statistical test/model chosen
Application of Test/Model	15	Correct execution of regression/test
Accuracy of Results	10	Correct interpretation of p-values, coefficients, error metrics
Insights & Relevance	10	Practical meaning of findings
Documentation	5	Clear report with visuals/tables

---

### CA 3: Final Report & Presentation (50 Marks)

Criterion	Marks	Description
Report Structure	10	Clear sections: Intro, Methodology, Results, Conclusion
Depth of Analysis	15	Integration of descriptive + inferential methods
Interpretation & Recommendations	10	Practical insights, critical thinking
Presentation Skills	10	Confidence, clarity, visualization, Q&A handling
Formatting & References	5	Proper citations, clean formatting

---

### Suggested Topic Distribution

- **CA 1:** Basic topics → descriptive statistics, probability, EDA.

- **CA 2:** *Intermediate topics* → hypothesis testing, regression, correlation, chi-square, ANOVA.
- **CA 3:** *Final integrated application* → advanced analysis (logistic regression, time series, clustering, or real-world case study with report + presentation).

## Workplan for Statistics Project (3 CA's)

---

### Week 1–2 → Orientation & Topic Finalization

- Introduction to project guidelines, datasets, and rubrics.
- Students (individually/groups) choose topics.
- Faculty approval of topics to ensure diversity.

---

### CA 1: Descriptive Statistics & EDA (50 Marks)



#### Tasks:

- Data collection/selection.
- Data cleaning (handling missing values, outliers).
- Descriptive statistics (mean, median, mode, variance, skewness, kurtosis).
- Graphical analysis (histograms, boxplots, scatter plots, bar/pie charts).

**Submission:** Report (tables + graphs + interpretation).

**Evaluation:** Rubrics for CA1 (50 Marks).

---

### CA 2: Statistical Modelling & Inference (50 Marks)



#### Tasks:

- Apply inferential statistics and models:
  - Hypothesis testing (t-test, chi-square, ANOVA).
  - Correlation & regression.

- Probability-based models.
- Interpret results with real-world meaning.

**Submission:** Short report with methodology, test results, and insights.

**Evaluation:** Rubrics for CA2 (50 Marks).

---

### **CA 3: Final Report & Presentation (50 Marks)**

#### **Week 10–12 (Final Assessment)**

##### **Tasks:**

- Compile work from CA1 + CA2.
- Extend with advanced/statistical modeling (logistic regression, clustering, or forecasting if dataset allows).
- Prepare **final comprehensive report** (Intro, Methodology, Results, Interpretation, Conclusion, References).
- Deliver **presentation/viva** (10–15 minutes).

**Submission:** Report + Presentation slides.

**Evaluation:** Rubrics for CA3 (50 Marks).

## (EXAMPLE)

### Workplan: Statistical Study of Mobile Phone Usage Patterns

---

#### Phase 1: Topic Introduction & Planning (Week 1–2)

##### Tasks:

- Define objectives (e.g., *to analyze screen time, app usage, and their impact on productivity/health*).
  - Decide data source:
    - Primary → student survey (Google Form).
    - Secondary → datasets from Kaggle (e.g., smartphone usage data).
  - Identify variables:
    - Daily screen time (hrs), number of apps used, most-used apps, internet data consumption, age, gender, sleep hours, productivity levels, etc.
  - Group/individual project allocation.
- 

#### CA 1: Descriptive Statistics & EDA (Week 3–4) – 50 Marks

##### Tasks:

- Collect dataset (minimum 50–100 responses if survey-based).
- Clean dataset (handle missing, group categorical variables).
- Perform **descriptive analysis**:
  - Central tendency (mean screen time, median daily calls, mode of app type).
  - Dispersion (variance/standard deviation of data usage).
  - Shape (skewness of screen time distribution).

- Visualization: histograms of screen time, pie chart of app usage categories, boxplot for daily calls.

**Submission:** Report with tables, charts, and interpretations.

##### Evaluation Rubric (50 Marks):

- Data collection & cleaning – 10

- Descriptive statistics – 15
  - Visualizations – 10
  - Insights – 10
  - Report quality – 5
- 

## **CA 2: Statistical Modelling & Inference (Week 6–8) – 50 Marks**

### **Tasks:**

- Apply **hypothesis testing**:
    - Example: *Is there a significant difference in mobile usage between male and female students?* (t-test/ANOVA).
    - Example: *Is there an association between mobile usage and sleep hours?* (Chi-square test).
  - **Correlation & regression**:
    - Relationship between screen time and academic performance/productivity.
    - Regression predicting internet data usage from screen time.
  - Interpret **p-values, confidence intervals, regression coefficients**.
- Submission:** Mini-report (methodology, statistical tests applied, results, conclusions).
- Evaluation Rubric (50 Marks):**
- Selection of methods – 10
  - Correct execution of tests – 15
  - Interpretation of results – 10
  - Practical relevance – 10
  - Documentation – 5
- 

## **CA 3: Final Report & Presentation (Week 10–12) – 50 Marks**

### **Tasks:**

- Combine **CA1 + CA2 analysis** into a comprehensive report.
- Add **advanced analysis**:

- Cluster students into groups (light, moderate, heavy users).
- Logistic regression: predict probability of “sleep disturbance = Yes/No” based on screen time.
- Prepare a **presentation (10–15 mins)** with:
  - Introduction, dataset, methodology, results, recommendations.
  - Charts/graphs + statistical evidence.
- Present findings in class.

**Evaluation Rubric (50 Marks):**

- Report structure – 10
- Depth of analysis (integration of CA1 + CA2 + advanced method) – 15
- Interpretation & recommendations – 10
- Presentation skills (clarity, confidence, visuals) – 10
- Formatting & references – 5

#### **Phase 4: Wrap-up (Week 13–14)**

- Reflection: *What insights did we gain about phone usage?*
- Faculty feedback on statistical rigor and real-world application.

#### **Summary Timeline**

Week	Activity	Deliverable	Marks
1–2	Topic planning, dataset design/collection	Dataset + objectives	–
3–4	<b>CA1: Descriptive &amp; EDA</b>	Report (stats + visuals)	50
6–8	<b>CA2: Modelling &amp; Inference</b>	Mini-report (tests + regression)	50
10–12	<b>CA3: Final Report + Presentation</b>	Comprehensive report + PPT	50
13–14	Reflection & feedback	–	–