

**LAB WORKBOOK**

Koneru Lakshmaiah Educational Foundation

**23IE02RF - RESEARCH METHODOLOGY, ETHICS SCIENTIFIC WRITING**



Koneru Lakshmaiah Educational Foundation

**LAB WORKBOOK**

|  |  |
| --- | --- |
| STUDENT NAME |  |
| REG.NO |  |
| YEAR |  |
| SEMESTER |  |
| SECTION |  |
| FACULTY |  |

Research Methodology

**Organization of the student lab workbook**

The laboratory framework includes a creative element but shifts the time-intensive aspects out- side of the Two-Hour closed laboratory period. Within this structure, each laboratory includes three parts: Prelab, In-lab, and Post-lab.

a **Pre-Lab**: The Prelab exercise is a homework assignment that links the lecture with the laboratory period - typically takes 2 hours to complete. The goal is to synthesize the infor- mation they learn in lecture with material from their textbook to produce a working piece of software. Prelab Students attending a two-hour closed laboratory are expected to make a good-faith effort to complete the Prelab exercise before coming to the lab. Their work need not be perfect, but their effort must be real (roughly 80 percent correct).

1. **In-Lab**: The In-lab section takes place during the actual laboratory period. The First hour of the laboratory period can be used to resolve any problems the students might have experienced in completing the Prelab exercises. The intent is to give constructive feedback so that students leave the lab with working Prelab software - a significant accomplishment on their part. During the second hour, students complete the In-lab exercise to reinforce the concepts learned in the Prelab. Students leave the lab having received feedback on their Prelab and In-lab work.
2. **Post-Lab**: The last phase of each laboratory is a homework assignment that is done following the laboratory period. In the Post-lab, students analyse the efficiency or utility of a given system call. Each Post-lab exercise should take roughly 120 minutes to complete.

Koneru Lakshmaiah Educational Foundation

3

##### 2024-25 ODD SEMESTER TUTORIAL CONTINUOUS EVALUATION

Koneru Lakshmaiah Educational Foundation

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SL.NO** | **DATE** | **Experiment TITLE** | **PRE-Lab (10M)** | **IN-Lab (20M)** | **POST-LabL (15M)** | **VIVA- VOCE (5M)** | **TOTAL (50M)** | **FACULTY SIGNA- TURE** |
| 1 |  | Identification of the research problem |  |  |  |  |  |  |
| 2 |  | Preparation of research summary |  |  |  |  |  |  |
| 3 |  | Preparation of Questionnaire for qualitative  research / collection of data for quantitiative  research |  |  |  |  |  |  |
| 4 |  | Preprocessing of data |  |  |  |  |  |  |
| 5 |  | Identification of features for data analysis |  |  |  |  |  |  |
| 6 |  | Comparison of metrics using various  algorithms |  |  |  |  |  |  |
| 7 |  | Preparation of articles |  |  |  |  |  |  |
| 8 |  | Preparation of Book chapters |  |  |  |  |  |  |
| 9 |  | Bullets and numbering |  |  |  |  |  |  |
| 10 |  | Placing figures in article |  |  |  |  |  |  |
| 11 |  | Tables in latex |  |  |  |  |  |  |
| 12 |  | Preparation of journal paper |  |  |  |  |  |  |
| 13 |  | Ethical issues related to publishing plagiarism  and self plagiarism |  |  |  |  |  |  |
| 14 |  | Preparing of poster |  |  |  |  |  |  |
| 15 |  | Preparing of presentation slides |  |  |  |  |  |  |

23IE02RF - RESEARCH METHODOLOGY, ETHICS SCIENTIFIC WRITING

#### IRD Team

Koneru Lakshmaiah Educational Foundation

Research Methodology

Koneru Lakshmaiah Educational Foundation

**Koneru Lakshmaiah Educational Foundation**

**Contents**

###### Research problem 5

###### Aim/Objective 5

###### Description: 5

###### Steps in Identifying a Research Problem 5

###### Materials and database needed: 6

###### Pre-Lab Questions : 7

###### In-Lab Activities : 8

###### Introduction and Discussion (15 minutes): 8

###### Literature Review Exercise (30 minutes): 8

###### Identifying Gaps and Formulating Problems (30 minutes): 9

###### Presentation and Feedback (30 minutes): 9

###### Individual Reflection (15 minutes): 9

###### Post-Lab Questions : 9

* 1. Viva-voce 10

###### Literature Survey 15

* 1. Pre-Lab 15
  2. In-Lab 16
  3. [Post-Lab](#_bookmark0) 18

###### [Data Collection / Questionnaire](#_bookmark1) 23

###### [Pre-Lab Questions :](#_bookmark1) 23

###### [In-Lab Activities :](#_bookmark2) 24

###### [Introduction and Discussion (10 minutes):](#_bookmark2) 24

###### [Identifying Gaps and Formulating research Problem (30 minutes):](#_bookmark2) . 24

###### [Research question Framing Exercise (30 minutes):](#_bookmark2) 24

###### [Presentation and Feedback (30 minutes):](#_bookmark3) 25

###### [Post-Lab Questions :](#_bookmark3) 25

* 1. [Viva-voce](#_bookmark3) 25

###### [Data Preprocessing](#_bookmark4) 29

* 1. [Pre-Lab](#_bookmark4) 29
  2. [In-Lab](#_bookmark5) 31
  3. [Post-Lab](#_bookmark6) 33

###### [Dimensionality reduction](#_bookmark7) 35

* 1. [Pre-Lab](#_bookmark7) 35
  2. [In-Lab](#_bookmark8) 36
  3. [Post-Lab](#_bookmark9) 39

Research Methodology CONTENTS

1. [Article Preparation](#_bookmark10) 45
2. [Book Preparation](#_bookmark11) 47
3. [Bullets-numbering](#_bookmark12) 49
   1. [Introduction to bullets](#_bookmark12) 49
   2. [Introduction to numbering 50](#_TOC_250006)
   3. [Excercise 51](#_TOC_250005)
4. Figures 53
   1. [Figures 53](#_TOC_250004)
   2. [Multiple figures 54](#_TOC_250003)
   3. [Excercise 55](#_TOC_250002)
5. [Tables](#_bookmark13) 59
   1. [Table in LATEX](#_bookmark13) 59
   2. [Cells spanning multiple rows or multiple columns](#_bookmark14) 60
   3. [Resizing the table](#_bookmark15) 61
   4. [Producing landscape tables](#_bookmark15) 61
   5. [Excercise](#_bookmark16) 63
6. [Book/Thesis-Preparation](#_bookmark17) 65
   1. [Aim/Objectives](#_bookmark17) 65
   2. [Building the chapters in the book](#_bookmark17) 65
7. [Book/Thesis-Preparation](#_bookmark18) 69

Koneru Lakshmaiah Educational Foundation

1. [Ethical Issues](#_bookmark19) 73
   1. [DEFINITION OF PLAGIARISM](#_bookmark19) 73
   2. [TYPES OF PLAGIARISM](#_bookmark20) 74
   3. [Preventing Plagiarism – Student Resources](#_bookmark21) 75
   4. [HOW TO AVOID PLAGIARISM](#_bookmark22) 77
      1. [BEST PRACTICES TO AVOID PLAGIARISM](#_bookmark22) 77
   5. [What is Citation?](#_bookmark23) 79
      1. [Doesn’t citing sources make my work seem less original?](#_bookmark23) 79
      2. [How do I cite sources?](#_bookmark23) 79
2. [Poster](#_bookmark24) 81
   1. [Introduction Poster](#_bookmark24) 81
3. Slides 87
   1. [Presentation 87](#_TOC_250001)
   2. [Creation of slides using LATEXBeamer class 88](#_TOC_250000)

**Experiment 4**

**Data Preprocessing**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

###### Learning outcomes:

* Understanding the data relevancy
* Able to identify the missing values, outliers, inappropriate values and prepare the data set for further analysis
  1. **Pre-Lab**
     1. Why data are dirty?

Koneru Lakshmaiah Educational Foundation

1. Data can be considered "dirty" due to various reasons, such as errors, inconsistencies, or missing values. These issues often arise from human input errors, system glitches, or data integration from multiple sources with different formats. Incomplete or inaccurate data can result from problems like incorrect data entry, duplicated records, or changes in data collection methods over time. Additionally, data may become outdated, irrelevant, or biased, further complicating its reliability. These imperfections make it necessary to clean and preprocess the data before analysis to ensure valid and accurate results.
   * 1. What is data preprocessing? Why is it important in research?
2. Data preprocessing is the process of cleaning, transforming, and organizing raw data into a suitable format for analysis. It involves tasks like handling missing values, removing duplicates, correcting errors, and standardizing data. Preprocessing also includes normalizing or scaling data and encoding categorical variables. This step is important in research because it ensures the quality, consistency, and reliability of the data, which directly affects the accuracy of the analysis and conclusions. Well-prepared data leads to more meaningful results, reduces biases, and enhances the overall efficiency of the research process.

Research Methodology EXPERIMENT 4. DATA PREPROCESSING

* + 1. What are some common problems that occur during data processing? How can they be fixed?

1. Common problems during data processing include missing values, duplicate records, inconsistent formats, and outliers.

* Missing values can be fixed by either imputing missing data using mean, median, or mode, or by removing the incomplete records if they are not critical.
* Duplicate records can be resolved by identifying and removing duplicates, ensuring that each record is unique.
* Inconsistent formats can be addressed by standardizing data formats, such as date formats or currency symbols, to ensure uniformity.
* Outliers can be handled by using statistical methods to detect and either remove or adjust extreme values, depending on their relevance to the analysis.

These issues must be addressed during data preprocessing to improve the quality and accuracy of the data for analysis.

* + 1. How do you handle the missing data?

1. Handling missing data can be done in several ways, depending on the extent and nature of the missing values. Common approaches include:

1. Imputation : Filling in missing values with a calculated value, such as the mean, median, or mode of the column, or using more advanced techniques like regression or K-nearest neighbors (KNN) imputation.

2. Deletion : Removing rows with missing values if the amount of missing data is small and not likely to affect the analysis. This is commonly used when the missing data is random and does not represent a significant portion of the dataset.

3. Interpolation : Estimating missing values based on available data points, particularly useful for time series data, where missing values can be filled using neighboring values.

4. Predictive Models : Using machine learning algorithms to predict and fill missing values based on patterns found in other features.

The chosen method depends on the context, the amount of missing data, and its potential impact on the analysis.

* + 1. What is the difference between missing value treatment and outliers treatment?

Koneru Lakshmaiah Educational Foundation

1. The main difference between missing value treatment and outlier treatment lies in the nature of the data issues they address and how they are handled:

* Missing value treatment focuses on handling absent data. Missing values can be addressed by imputation (replacing with mean, median, mode, etc.), deletion (removing rows or columns with missing data), or interpolation (estimating values based on other data points). The goal is to fill gaps in the dataset so that analysis can proceed without losing important information.
* Outlier treatment deals with data points that are significantly different from the rest of the data. Outliers can distort statistical analysis and models. They can be handled by removing them, transforming the data (e.g., using log transformations), or capping values within a defined range (winsorization). The goal is to address extreme values that may skew the results or analysis.

In summary, missing value treatment aims to address gaps in the data, while outlier treatment focuses on managing extreme or abnormal values that deviate from the expected range.

Research Methodology EXPERIMENT 4. DATA PREPROCESSING

Koneru Lakshmaiah Educational Foundation

* 1. **In-Lab**

Analyze the the following Data set and interpret the dataset.

* + 1. [Diabetes Dataset](https://www.kaggle.com/datasets/mathchi/diabetes-data-set/)

Identify the missing values, any inaccuracies, outliers present in the dataset. Do the neces- sary transformation such as normalization, encoding and feature scaling

1. Highlighting the distinctions between missing value treatment and outlier treatment, both are essential preprocessing steps in data cleaning, but each targets a specific type of data issue:

* **Missing Value Treatment**: This involves dealing with absent or undefined values within a dataset, which can occur due to data entry errors, data unavailability, or technical issues during data collection. Strategies include:
* **Imputation**: Replacing missing values with statistical measures like the mean, median, or mode, or using more sophisticated methods like k-nearest neighbors or predictive modeling for estimation.
* **Deletion**: Removing rows or columns with a significant proportion of missing data, especially when imputation might introduce bias or when missing values are non-random.
* **Interpolation**: Estimating missing values by using patterns or trends in surrounding data, especially useful for time-series data.
* **Outlier Treatment**: This addresses values that are drastically different from the majority of data points and could lead to skewed results or erroneous conclusions. Common approaches include:
* **Removal**: Eliminating outliers, often based on a defined threshold or statistical measure (e.g., values beyond 1.5 times the interquartile range).
* **Transformation**: Applying mathematical transformations, like log or square root transformations, to reduce the impact of extreme values.
* **Winsorization**: Capping extreme values to fall within a defined percentile range, which limits their influence without discarding them outright.

Both treatments are necessary for improving data quality, but while missing value treatment fills in incomplete information, outlier treatment minimizes the influence of atypical data points on the analysis.

Research Methodology EXPERIMENT 4. DATA PREPROCESSING

Koneru Lakshmaiah Educational Foundation

* + 1. [Movies data set](https://www.kaggle.com/datasets/bharatnatrayn/movies-dataset-for-feature-extracion-prediction/data?select=movies.csv) Identify the missing values, any inaccuracies, outliers present in the dataset. Do the necessary transformation such as normalization, encoding and feature scaling

**ANSWER :**

To handle the **Movies dataset** effectively, let’s break down each preprocessing step to ensure the data is ready for analysis or modeling:

**1. Identify and Handle Missing Values**

* **Identify Missing Values**: Check for missing values in each column, either using a .isnull().sum() method in Python (e.g., pandas) or visually inspecting a summary table.
* **Handle Missing Values**:
  + **Impute**: For numerical columns, fill missing values with statistical values like mean or median, or use more advanced techniques such as k-nearest neighbors for estimation.
  + **Categorical Variables**: Fill missing categorical values with the most frequent category (mode), or create a separate “Unknown” category.
  + **Drop Missing Values**: If a column has an excessive number of missing values (e.g., >70%), consider dropping it if it does not significantly impact the analysis.

**2. Identify and Address Inaccuracies**

* **Inconsistencies in Data**: Look for inaccuracies such as incorrect data types, spelling errors in categorical values (like genres or director names), or misplaced values (e.g., extremely high/low values that don’t match the context).
* **Correct Inaccuracies**: Standardize text values, convert data types if necessary, and review any logically improbable values (like negative durations) for correction or removal.

**3. Outlier Detection and Treatment**

* **Detect Outliers**: Use visualization techniques such as box plots or scatter plots, or statistical methods (like z-scores) to identify extreme values in numerical columns like revenue, ratings, or runtime.
* **Handle Outliers**:
  + **Removal**: Remove outliers if they seem to be data errors or highly unlikely values.
  + **Transformation**: Apply transformations (e.g., log or square root) to reduce the influence of extreme values, especially for skewed distributions.
  + **Winsorization**: Cap outlier values to fall within a specified range, particularly for features where outliers might be valid but overly influential.

**4. Transformations (Normalization, Encoding, and Feature Scaling)**

* **Normalization**: For features like revenue or budget that may span a large range, normalize them (e.g., min-max scaling to a range of [0,1]) to improve comparability.
* **Encoding Categorical Variables**:
  + **One-Hot Encoding**: For categorical variables with nominal values, such as genre, director, or language, apply one-hot encoding to convert categories into binary columns.
  + **Ordinal Encoding**: If any categorical data has a defined order (like movie ratings), use ordinal encoding to maintain that order.
* **Feature Scaling**:
  + **Standardization**: For numerical features with different scales, use standardization (mean of 0, standard deviation of 1) to make features comparable and improve the performance of distance-based models.
  + **Robust Scaling**: For features with extreme outliers, use robust scaling to reduce the effect of outliers by scaling based on the interquartile range.

Research Methodology EXPERIMENT 4. DATA PREPROCESSING

Koneru Lakshmaiah Educational Foundation

After completing these steps, the **Movies dataset** will be cleaned and preprocessed, ready for analysis or further modeling steps. This process minimizes biases from missing values, inconsistencies, and outliers, while enhancing the dataset's usability.

* 1. **Post-Lab**

Implement a Python program to apply various data preprocessing techniques on the following dataset.

[Dataset Link](https://catalog.data.gov/dataset/electric-vehicle-population-data/resource/fa51be35-691f-45d2-9f3e-535877965e69) Identify the missing values, any inaccuracies, outliers present in the dataset. Do the necessary transformation such as normalization, encoding and feature scaling

**ANSWER :**

# Import necessary libraries

import pandas as pd

import numpy as np

from sklearn.impute import SimpleImputer

from sklearn.preprocessing import StandardScaler, MinMaxScaler, OneHotEncoder

from sklearn.compose import ColumnTransformer

from sklearn.pipeline import Pipeline

import matplotlib.pyplot as plt

import seaborn as sns

# Load the dataset

df = pd.read\_csv("dataset.csv")

# Step 1: Identifying Missing Values

print("Missing values in each column:")

print(df.isnull().sum())

# Step 2: Handling Missing Values

# For numerical columns, fill missing values with the mean

numeric\_cols = df.select\_dtypes(include=['float64', 'int64']).columns

imputer\_num = SimpleImputer(strategy='mean')

df[numeric\_cols] = imputer\_num.fit\_transform(df[numeric\_cols])

# For categorical columns, fill missing values with the most frequent value

categorical\_cols = df.select\_dtypes(include=['object']).columns

imputer\_cat = SimpleImputer(strategy='most\_frequent')

df[categorical\_cols] = imputer\_cat.fit\_transform(df[categorical\_cols])

# Step 3: Identifying Outliers

# Plot boxplots to visualize outliers in numerical features

plt.figure(figsize=(10, 6))

for i, col in enumerate(numeric\_cols, 1):

plt.subplot(2, (len(numeric\_cols) + 1) // 2, i)

sns.boxplot(y=df[col])

plt.title(f'Boxplot of {col}')

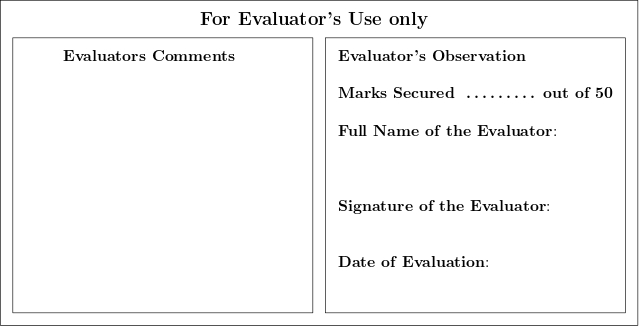
plt.tight\_layout()

plt.show()

# Step 4: Handling Outliers

# Option 1: Remove outliers based on z-score threshold (e.g., |z| > 3)

from scipy.stats import zscore

Research Methodology EXPERIMENT 4. DATA PREPROCESSING

Koneru Lakshmaiah Educational Foundation

z\_scores = np.abs(zscore(df[numeric\_cols]))

outliers = (z\_scores > 3)

df = df[(~outliers).all(axis=1)]

# Step 5: Transformations - Encoding, Normalization, and Scaling

# One-hot encode categorical variables

df = pd.get\_dummies(df, columns=categorical\_cols, drop\_first=True)

# Standardize numerical features

scaler = StandardScaler()

df[numeric\_cols] = scaler.fit\_transform(df[numeric\_cols])

# Alternatively, use Min-Max Scaling if preferred

# min\_max\_scaler = MinMaxScaler()

# df[numeric\_cols] = min\_max\_scaler.fit\_transform(df[numeric\_cols])

# Check final processed data

print("Sample of the processed dataset:")

print(df.head())

**Experiment 5**

**Identification of features for data analysis and Dimensionality reduction**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

###### Learning outcomes:

* Understanding the features
* Dimensionality reduction
  1. **Pre-Lab**

Koneru Lakshmaiah Educational Foundation

* + 1. What is dimensionality reduction?

1. Dimensionality reduction is a technique to reduce the number of features in a dataset while preserving essential information. It simplifies data, reduces computational cost, and helps prevent overfitting by removing noise and eliminating redundant features, which can improve model accuracy. It’s especially useful in high-dimensional datasets, making patterns and clusters easier to identify and interpret. Common methods include Principal Component Analysis (PCA) and t-SNE for visualization.
   * 1. Why is dimensionality reduction important?
2. Dimensionality reduction is important because it simplifies high-dimensional data, making analysis more efficient and less prone to overfitting. By reducing the number of features, it decreases computational costs and helps models focus on the most relevant information, leading to potentially higher accuracy. It also aids in data visualization, allowing patterns, clusters, or relationships to become more apparent in a reduced feature space. This process is especially valuable in handling noisy or redundant data, as it improves model interpretability and effectiveness.
   * 1. What are some common techniques used for dimensionality reduction?
3. Some common techniques for dimensionality reduction include:
4. **Principal Component Analysis (PCA)**: Transforms data into a smaller set of uncorrelated variables (principal components) that capture the most variance in the data.
5. **t-Distributed Stochastic Neighbor Embedding (t-SNE)**: Non-linear technique used mainly for visualizing high-dimensional data in 2D or 3D by preserving local structure.

Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

1. **Linear Discriminant Analysis (LDA)**: Similar to PCA, but it also maximizes class separability, making it useful for classification problems.
   * 1. **Autoencoders**: Neural networks used to learn an efficient, lower-dimensional representation of the input data through encoding and decoding layers.
2. How does Principal Component Analysis (PCA) work?

PCA works by transforming data into new axes called \*\*principal components\*\* that capture the most variance. It starts by standardizing the data, then computes the covariance matrix to understand feature relationships. Eigenvalues and eigenvectors are calculated to identify the principal components, and finally, data is projected onto these components, reducing dimensionality while retaining important information.

* + 1. What is the difference between PCA and LDA?

PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis) both reduce dimensionality but serve different purposes. PCA focuses on maximizing variance and is unsupervised, while LDA maximizes class separability and is supervised, using class labels to find the best feature directions.

Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

Koneru Lakshmaiah Educational Foundation

* 1. **In-Lab**
     1. Use the [Principle component Analysis,](https://www.datacamp.com/tutorial/principal-component-analysis-in-python) analyse Breast Cancer data set is a real-valued mul- tivariate data that consists of two classes, where each class signifies whether a patient has breast cancer or not [refer the link for the study h[ttps://www.datacamp.com/tutorial/principal-](http://www.datacamp.com/tutorial/principal-) component-analysis-in-python]. The two categories are: malignant and benign. Download

the dataset using the [link](https://www.cs.toronto.edu/~kriz/cifar.html) [h[ttps://www.cs.toronto.edu/](http://www.cs.toronto.edu/) kriz/cifar.html]

Koneru Lakshmaiah Educational Foundation

**ANSWER :**

To perform Principal Component Analysis (PCA) on the Breast Cancer dataset using Python, follow the steps below. The dataset you're referring to seems to be the **Breast Cancer Wisconsin dataset** (often available from other sources too). Here, I’ll guide you through the steps using scikit-learn to load the data, apply PCA, and analyze the results.

**Steps to Perform PCA on the Breast Cancer Dataset**

1. **Install necessary libraries** (if not already installed):

bash

Copy code

pip install pandas scikit-learn matplotlib seaborn

1. **Download and load the dataset**: You can use sklearn.datasets.load\_breast\_cancer() to load the Breast Cancer dataset. If you have a CSV file, use pandas.read\_csv() to load it.
2. **Apply PCA**: Perform PCA on the dataset to reduce the dimensionality and visualize the results.

Full Python Code Example

# Import necessary libraries

import pandas as pd

import numpy as np

from sklearn.decomposition import PCA

from sklearn.preprocessing import StandardScaler

from sklearn.datasets import load\_breast\_cancer

import matplotlib.pyplot as plt

import seaborn as sns

# Load the dataset from sklearn

data = load\_breast\_cancer()

X = data.data # Features

y = data.target # Labels (0 = malignant, 1 = benign)

Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

Koneru Lakshmaiah Educational Foundation

# Standardize the features

scaler = StandardScaler()

X\_scaled = scaler.fit\_transform(X)

# Apply PCA to reduce to 2 components

pca = PCA(n\_components=2)

X\_pca = pca.fit\_transform(X\_scaled)

# Create a DataFrame for the PCA results

df\_pca = pd.DataFrame(data=X\_pca, columns=['PC1', 'PC2'])

df\_pca['label'] = y

# Plot the PCA results

plt.figure(figsize=(8,6))

sns.scatterplot(data=df\_pca, x='PC1', y='PC2', hue='label', palette={0: 'red', 1: 'blue'}, s=60)

plt.title('PCA of Breast Cancer Dataset')

plt.xlabel('Principal Component 1')

plt.ylabel('Principal Component 2')

plt.legend(title='Class', labels=['Malignant', 'Benign'])

plt.show()

# Explained variance

explained\_variance = pca.explained\_variance\_ratio\_

print(f"Explained variance by each component: {explained\_variance}")

Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

* 1. **Post-Lab**
     1. What is factor analysis?

Factor analysis is a statistical method used to identify underlying relationships between variables by grouping them into factors that explain the variance in the data.

* + 1. What are the types of factor analysis?

The two main types of factor analysis are:

1. **Exploratory Factor Analysis (EFA)**: Identifies underlying factors without a predefined structure.
2. **Confirmatory Factor Analysis (CFA)**: Tests a hypothesized factor structure based on existing theory.

* + 1. What are eigenvalues in factor analysis?

Eigenvalues in factor analysis represent the amount of variance explained by each factor. A higher eigenvalue indicates a more significant factor.

* + 1. How is factor loading interpreted?

Koneru Lakshmaiah Educational Foundation

Factor loading indicates the correlation between a variable and a factor. Higher loadings (closer to 1 or -1) suggest a stronger relationship, while lower loadings (closer to 0) indicate a weaker relationship.

* + 1. Differentiate between the PCA and Factor analysis.

PCA focuses on reducing dimensionality by maximizing variance in the data, while **Factor Analysis** identifies underlying factors that explain correlations between variables. PCA is more about data compression, and Factor Analysis aims at uncovering latent structures.

Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

Koneru Lakshmaiah Educational Foundation

1. Use the [Factor Analysis,](https://www.datacamp.com/tutorial/introduction-factor-analysis) analyse factor BFI (dataset based on personality assessment project), which were collected using a 6 point response scale: 1 Very Inaccurate, 2 Moderately Inac- curate, 3 Slightly Inaccurate 4 Slightly Accurate, 5 Moderately Accurate, and 6 Very Accu- rate. You can also download this dataset from the following the [link:](https://vincentarelbundock.github.io/Rdatasets/datasets.html)

###### Solution

To analyze the **BFI (Big Five Inventory) dataset** using **Factor Analysis**, follow these steps:

**1. Install Required Libraries**

Ensure you have the necessary libraries installed:

bash

Copy code

pip install pandas numpy scikit-learn matplotlib seaborn

**2. Import Libraries and Load the Dataset**

You’ll first need to download the dataset and load it into your environment. Assuming you have the dataset in CSV format, you can use pandas to load it.

python

Copy code

import pandas as pd

import numpy as np

from sklearn.decomposition import FactorAnalysis

import matplotlib.pyplot as plt

import seaborn as sns

# Load the BFI dataset (replace 'file\_path' with the actual path to the dataset)

df = pd.read\_csv('file\_path') # Replace with actual file path

**3. Preprocessing the Data**

Ensure the data is clean, with no missing values. If necessary, handle missing values by imputation or removal.

python

Copy code

# Check for missing values

df.isnull().sum()

# Handle missing values (e.g., impute with the mean)

df.fillna(df.mean(), inplace=True)

Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

Koneru Lakshmaiah Educational Foundation

**4. Standardize the Data**

Factor Analysis works better when the data is standardized.

python

Copy code

from sklearn.preprocessing import StandardScaler

# Standardize the data

scaler = StandardScaler()

X\_scaled = scaler.fit\_transform(df)

**5. Apply Factor Analysis**

Choose the number of factors to extract. A common approach is to use an eigenvalue greater than 1 rule or trial and error to determine the number of factors.

python

Copy code

# Apply Factor Analysis to extract 5 factors (one for each Big Five personality trait)

fa = FactorAnalysis(n\_components=5)

X\_fa = fa.fit\_transform(X\_scaled)

**6. Factor Loadings and Interpretation**

The factor loadings tell you how strongly each variable relates to each factor. You can check the factor loadings and interpret them.

python

Copy code

# Get the factor loadings

loadings = pd.DataFrame(fa.components\_, columns=df.columns)

# Display the factor loadings

print(loadings)

**7. Visualize the Results**

You can visualize the factor loadings or the transformed data.

python

Copy code

# Heatmap of factor loadings

plt.figure(figsize=(10, 8))

sns.heatmap(loadings, annot=True, cmap='coolwarm', fmt='.2f')

plt.title('Factor Loadings from Factor Analysis')

plt.show()

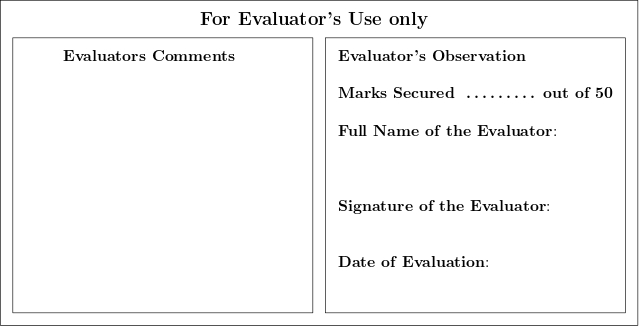
Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

Koneru Lakshmaiah Educational Foundation

**8. Analyze the Results**

* Each factor will represent a combination of the original variables, which can be interpreted based on their loadings.
* Typically, each factor corresponds to one of the **Big Five Personality Traits** (e.g., Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism), depending on how the items in the dataset are structured.

This process will help you identify latent factors in the BFI dataset and interpret them in terms of the Big Five personality traits.

Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

Koneru Lakshmaiah Educational Foundation

Research Methodology EXPERIMENT 5. DIMENSIONALITY REDUCTION

Koneru Lakshmaiah Educational Foundation

**Experiment 6**

**Preparation of simple article**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

\begin{document}

%generate the title

\maketitle

%insert the table of contents

\tableofcontents

% \part{Part - I - KLEF}

\chapter{Introduction}

\section{Introduction}

\label{sec:IRD}

Koneru Lakshmaiah Educational Foundation

KLEF Deemed to be university was established in 1980-81, as KL College of Engineering, which

The institutes remains indebted to leadership of our late co-founder Er.Koneru Lakshmaiah, an

\section{Description on CSE Department}

The B.Tech COMPUTER SCIENCE \& INFORMATION TECHNOLOGY (CSIT) is started in the year 2020 whic

This four-year undergraduate program is a blend of engineering and tool-based skill-driven

\begin{comment}

\par In section \ref{sec:IRD}, gives a detail account on KLEF IRD department

Globally, Information Technology (IT) skills continue to be in high demand from businesses in

\end{comment}

\section{Description on AIDS Department}

The program of Artificial Intelligence \& Data Science started in the year 2020. There are mo

The faculty and students of the program pursue their research in various fields like machine

The program offers programs like B.Tech (AI\&DS), B.Tech (AI\&DS) with Hons, B.Tech (AI\&DS)

\section{Introduction}

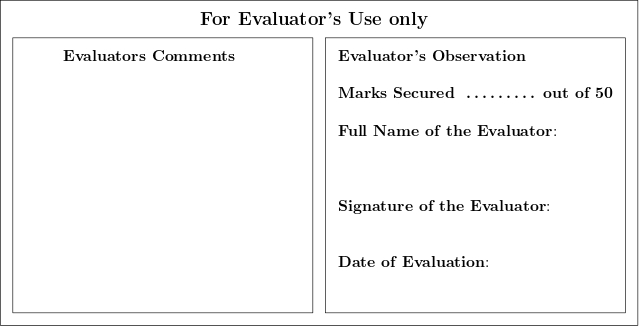
The Department of CSE has started in the year 1980 and has grown up exponentially in size and

\end{document}

Research Methodology EXPERIMENT 6. ARTICLE PREPARATION

[a4paper,11pt]book [bottom=1in,top=1in,left=1in,right=1in]geometry graphicx amsmath tabu- larx fancyhdr

Koneru Lakshmaiah Educational Foundation



**Experiment 7**

**Preparation of Book Chapters**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

\begin{document}

%generate the title

\maketitle

%insert the table of contents

\tableofcontents

\part{Part - I - KLEF}

\chapter{KLEF}

\section{Introduction}

Koneru Lakshmaiah Educational Foundation

KLEF Deemed to be university was established in 1980-81, as KL College of Engineering, which

The institutes remains indebted to leadership of our late co-founder Er.Koneru Lakshmaiah, a

\chapter{About CSE}

\section{Introduction}

The Department of CSE has started in the year 1980 and has grown up exponentially in size and

% \chapter{About CSIT}

\section{Introduction}

The B.Tech COMPUTER SCIENCE \& INFORMATION TECHNOLOGY (CSIT) is started in the year 2020 whic

This four-year undergraduate program is a blend of engineering and tool-based skill-driven

Globally, Information Technology (IT) skills continue to be in high demand from businesses in

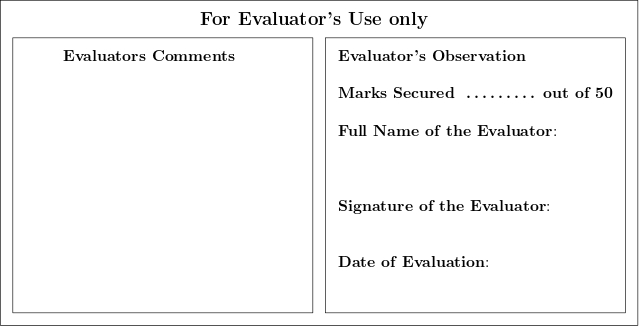
% \chapter{About AIDS}

\section{Introduction}

The program of Artificial Intelligence \& Data Science started in the year 2020. There are mo

The faculty and students of the program pursue their research in various fields like machine The program offers programs like B.Tech (AI\&DS), B.Tech (AI\&DS) with Hons, B.Tech (AI\&DS)

\end{document}

Research Methodology EXPERIMENT 7. BOOK PREPARATION

Koneru Lakshmaiah Educational Foundation

**Experiment 8**

**Bullets and Numbering**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

# Introduction to bullets

Bullet lists are commonly used in documents to present information in a concise and Unordered organized manner. The itemize environment is the most commonly used method for creating bul- let lists in LaTeX.

It automatically generates bullet points for each item in the list. To create a bullet list using the itemize environment, the following syntax is implemented. The output of the code will be like

Koneru Lakshmaiah Educational Foundation

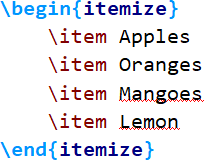


Figure 8.1: The example of Bulleted text in itemize environment

* + - Apples
    - Oranges
      * shimla orange
      * imported oranges
    - Mangoes
    - Lemon

Each item in the list is preceded by the \item command which causes LaTeX to automatically insert bullet points for each item.

Similarly, we can have nested bulleted items too. Example is given below:

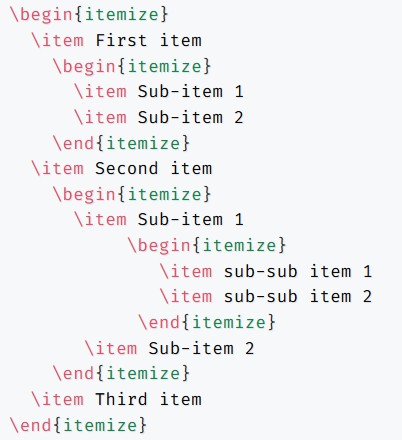
Research Methodology EXPERIMENT 8. BULLETS-NUMBERING

Figure 8.2: The example of enumerated text in enumerate environment

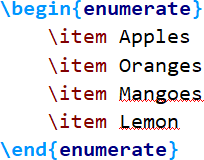
# Introduction to numbering

Lakshmaiah Educational Foundation

In some cases, you may need to create a numbered list to represent items in a more orderly fash- ion. This is useful for showing steps or sequences. You can create a numbered list with LATEX bul-

let points with the same code we used before, except with \begin{enumerate} and \end{enumerate}

around the list items instead of \begin{itemize} and \end{itemize}. The output of the above

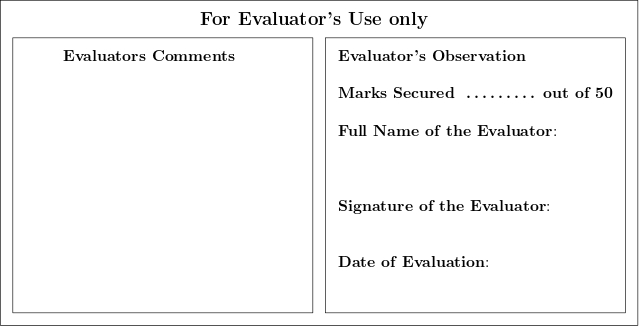


code will be like the following

1. Apples

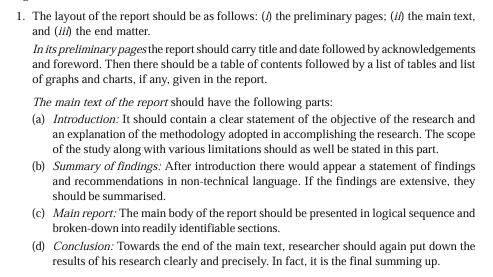
Koneru

1. Oranges
2. Mangoes
3. Lemon

Research Methodology EXPERIMENT 8. BULLETS-NUMBERING

Koneru Lakshmaiah Educational Foundation

# Excercise

The students are instructed to do the following excercise

Research Methodology EXPERIMENT 8. BULLETS-NUMBERING

Koneru Lakshmaiah Educational Foundation

**Experiment 9**

**Placing Figures in article**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

# Figures

Foundation

It’s necessary to add pictures to your documents. Using LATEXall pictures will be indexed auto- matically and tagged with successive numbers when using the figure environment and the graph- icx package. The figure environment is given as the following code:

Koneru Lakshmaiah Educational



Figure 9.1: A boat.

\begin{figure}

\includegraphics[width=\linewidth]{RMRECORD/ex4-FIGURES/boat.png}

\caption{A boat.}

\label{fig:boat1}

\end{figure}

The reference to the figure can be given in the running text as follows:

Research Methodology EXPERIMENT 9. FIGURES

Figure \ref{fig:boat1} shows a boat.

The reference to the figure can be given in the running text as follows: Figure 9.1 shows a boat. Setting the float by adding [h!] behind the figure environment \begin tag will force the figure to be shown at the location in the document. Possible values are:

h (here) – same location t (top) – top of page

b (bottom) – bottom of page p (page) – on an extra page

! (override) – will force the specified location

However, I have only used the [h!] option so far. The float package (\usepackage{float}) al- lows to set the option to [H], which is even stricter than [h!].

# Multiple figures

Sometimes when writing a document, adding single images is not optimal, especially when the reader is supposed to compare several results or graphs. In such situations, it might be necessary to use a different environment, called subfigure. The subfigure environment allows you to place multiple images at a certain location next to each other and the usage is pretty straightforward. First you need to add the subcaption package to your preamble:

\documentclass{article}

Koneru Lakshmaiah Educational Foundation

\usepackage{graphicx}

\usepackage{subcaption}

\begin{document}

%...

\end{document}

Next, you need to add multiple subfigure environments within a figure environment.

\begin{figure}[h!]

\centering

\begin{subfigure}[b]{0.4\linewidth}

\includegraphics[width=\linewidth]{RMRECORD/ex4-FIGURES/coffee.jpg}

\caption{Coffee.}

\end{subfigure}

\begin{subfigure}[b]{0.4\linewidth}

\includegraphics[width=\linewidth]{RMRECORD/ex4-FIGURES/coffee.jpg}

\caption{More coffee.}

\end{subfigure}

\caption{The same cup of coffee. Two times.}

\label{fig:coffee}

\end{figure}

The output of the above code will be:

Research Methodology EXPERIMENT 9. FIGURES

Koneru Lakshmaiah Educational Foundation

# Excercise

1. Develope a code for getting the subfigures as per the image given below (you given free hand to use any figure)

Research Methodology EXPERIMENT 9. FIGURES

Figure 9.3: More coffee.

0.4

(a) b



Figure 9.2: Coffee.

0.4

Educational Foundation

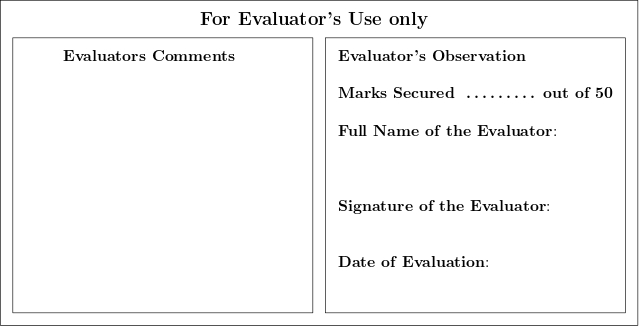
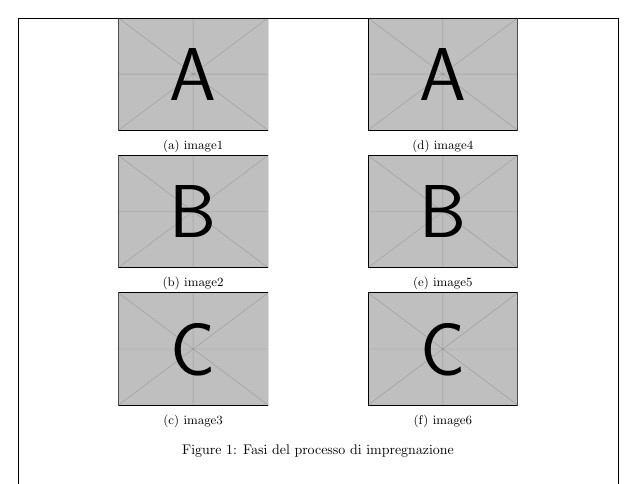
Lakshmaiah

(a) b

Koneru



56

Research Methodology EXPERIMENT 9. FIGURES

Educational Foundation

Koneru Lakshmaiah

Research Methodology EXPERIMENT 9. FIGURES

Koneru Lakshmaiah Educational Foundation

**Experiment 10**

**Tables in LATEX**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

# Table in LATEX

Foundation

Tables in LATEXcan be created through a combination of the table environment and the tabu- lar environment. The table environment part contains the caption and defines the float for our table, i.e. where in our document the table should be positioned and whether we want it to be displayed centered. The \caption and \label commands can be used in the same way as for pictures. The actual content of the table is contained within the tabular environment.

The tabular environment uses ampersands & as column seperators and newline symbols

Educational

as row seperators. The vertical lines separating the columns of our table (—) are passed as an argument to the tabular environment (e.g.

\begin{table}[htbp]

\begin{tabular}{l|c|r}

....

\end{tabular}

Lakshmaiah

\end{table}

) and the letters tell whether we want to align the content to the left (l), to the center (c) or to the right (r) for each column. There should be one letter for every column and a vertical line in between them or in front of them, if we want a vertical line to be shown in the table. Row seper- ators can be added with the \hline command.

Setting the float by adding [h!] behind the table environment \begin tag will force the table to be shown at the location in the document. Possible values are:

Koneru

h (here) – same location t (top) – top of page

b (bottom) – bottom of page p (page) – on an extra page

! (override) – will force the specified location

However, I have only used the [h!] option so far. The float package (\usepackage{float}) al- lows to set the option to [H], which is even stricter than [h!].

Research Methodology EXPERIMENT 10. TABLES

\begin{table}[!ht]

\centering

\begin{tabular}{|l|l|l|l|l|}

\hline

U/S & BGR & HYD & CON & Supply \\ \hline BZA & 15 & 60 & 35 & 15 \\ \hline

HYD & 45 & 30 & 60 & 30 \\ \hline

CHE & 30 & 90 & 20 & 20 \\ \hline Demand & 5 & 20 & 40 & \\ \hline

\end{tabular}

\end{table}

The above code produces the table (Table [10.1)](#_bookmark14)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| U/S | BGR | HYD | CON | Supply |
| BZA | 15 | 60 | 35 | 15 |
| HYD | 45 | 30 | 60 | 30 |
| CHE | 30 | 90 | 20 | 20 |
| Demand | 5 | 20 | 40 |  |

Table 10.1: The distance between various ciites Table [10.1,](#_bookmark14) gives the distance between different cities.

# Cells spanning multiple rows or multiple columns

Koneru Lakshmaiah Educational Foundation

Sometimes it’s necessary to make a row span several cells. For this purpose we can use the multi- row package, so the first thing we’re going to do is adding the required package to our preamble:

\usepackage{multirow}

In order for a cell to span multiple rows, we have to use the multirow command (refer Table [10.2).](#_bookmark15) This command accepts three parameters:

Research Methodology EXPERIMENT 10. TABLES

\multirow{NUMBER\_OF\_ROWS}{WIDTH}{CONTENT}

\begin{table}[!ht]

\centering

\begin{tabular}{|l|l|l|l|l|l|}

\hline

\multirow{2}{\*}{Store} & \multicolumn{4}{c|}{Retailshop}& \multirow{2}{\*}{Supply} \\ \cline{2

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| & | 1 | & | 2 | & | 3 | & | 4 | & | ~ | \\ | \hline |
| 1 | & | 3 | & | 5 | & | 7 | & | 6 | & | 80 | \\ \hline |
| 2 | & | 2 | & | 5 | & | 8 | & | 2 | & | 60 | \\ \hline |
| 3 | & | 3 | & | 6 | & | 9 | & | 2 | & | 70 | \\ \hline |

Demand & 40 & 60 & 50 & 60 & ~ \\ \hline

\end{tabular}

\end{table}

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Store | Retailshop | | | | Supply |
| 1 | 2 | 3 | 4 |
| 1 | 3 | 5 | 7 | 6 | 80 |
| 2 | 2 | 5 | 8 | 2 | 60 |
| 3 | 3 | 6 | 9 | 2 | 70 |
| Demand | 40 | 60 | 50 | 60 |  |

Table 10.2: Retail shop distances

Koneru Lakshmaiah Educational Foundation

# Resizing the table

To resize the table we need to add the required package in the preamble \usepackage{adjustbox}. We need use

\begin{table}

\adjustbox{max width=\linewidth}{%

\begin{tabular}{cols} content...

\end{tabular}

}

\end{table}

# Producing landscape tables

To produce the landscape tables, one needs to use landscape environment by adding the required package in the preamble \usepackage{pdflscape}

##### 2024-25 ODD SEMESTER TUTORIAL CONTINUOUS EVALUATION

Koneru Lakshmaiah Educational Foundation

Research Methodology

EXPERIMENT 10.

TABLES

62

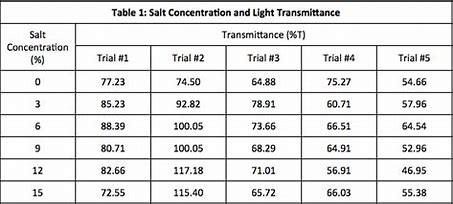
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SL.NO** | **DATE** | **TUTORIAL TITLE** | **PRE- TUTORIAL (10M)** | **IN- TUTORIAL (20M)** | **POST- TUTORIAL (15M)** | **VIVA- VOCE (5M)** | **TOTAL (50M)** | **FACULTY SIGNA- TURE** |
| 1 |  | Formulation of LPP and Graphical Method |  |  |  |  |  |  |
| 2 |  | Simplex Method and Principle of Duality |  |  |  |  |  |  |
| 3 |  | Transportation Problem |  |  |  |  |  |  |

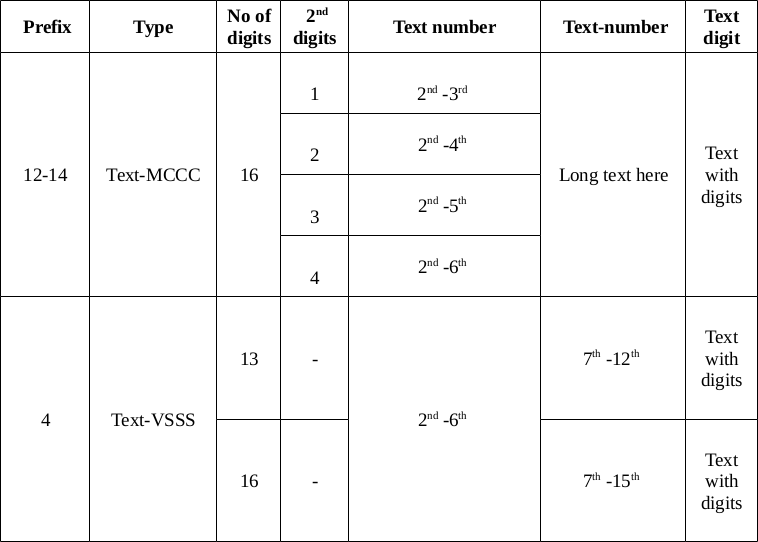
Research Methodology EXPERIMENT 10. TABLES

Koneru

# Excercise

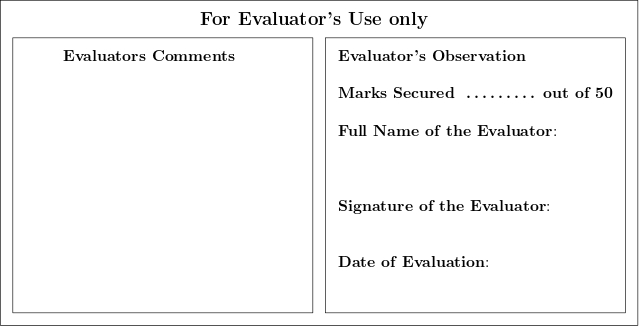
* + 1. Construct the latex code for the table given in the figure



* + 1. Construct the latex code for the table given in the figure

Lakshmaiah

Educational Foundation

Research Methodology EXPERIMENT 10. TABLES

Koneru Lakshmaiah Educational Foundation

**Experiment 11**

**Book/Thesis-Preparation**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

# Aim/Objectives

Foundation

The aim of this experiment is to understand how to build book or thesis:

* + - Creation of individual chapters
    - To include the created individual chapters into the mail file

Educational

# Building the chapters in the book

The steps to be followed to build the book:

1. Add all the preambles given in the individual chapters in the main LATEXfile
2. comment all the preambles given in the individual chapters.

Lakshmaiah

1. use the \input command to input the individual chapters The following code illustrates how to add the individual chapters:

\documentclass{book}

\usepackage{graphicx} % Required for inserting images

\usepackage{subcaption}

Koneru

\usepackage{adjustbox}

\usepackage{verbatim}

\usepackage{multirow}

\usepackage{caption}

\usepackage{pdfpages}

\title{Latex Workshop}

\author{kluirdteam }

\date{September 2024}

\begin{document}

\maketitle

\tableofcontents

Research Methodology EXPERIMENT 11. BOOK/THESIS-PREPARATION

Koneru Lakshmaiah Educational Foundation

\listoffigures

\listoftables

% \input{ex1-article/ex1-document-article}

% \input{ex2-book/ex2-document-book}

% \input{ex3-bullets\_numbering/ex3-document-bullets\_numbering}

% \input{ex4-FIGURES/ex4-Figures}

% \input{ex5-tables/ex5-tables}

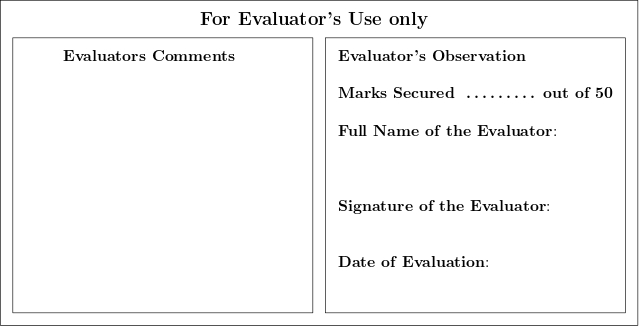
% \includepdf[pages=-]{Latex\_Workshop (1).pdf}

\end{document}

1. \maketitle - will create the title pages.
2. \tableofcontents - will create the table of contents mentioning the chapters, section along with the page numbers.
3. \listoffigures - will include the list the figures mentioned in the chapters.
4. \listoftables - will include the list of tables given in the chapters.

Research Methodology EXPERIMENT 11. BOOK/THESIS-PREPARATION

Koneru Lakshmaiah Educational Foundation



Research Methodology EXPERIMENT 11. BOOK/THESIS-PREPARATION

Koneru Lakshmaiah Educational Foundation

**Experiment 12**

**Preparation of Book / Thesis**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

###### Learning outcomes:

Foundation

* This experiment helps to prepare the book or Thesis
* Also modularity is introduced by means of chapters

\documentclass[a4paper,11pt]{book}

\usepackage[bottom=1in,top=1in,left=1in,right=1in]{geometry}

\usepackage{authblk}

\usepackage{amsmath}

Educational

\usepackage{hyperref}

\usepackage{adjustbox}

\usepackage{listings}

\usepackage{subfigure}

\usepackage{tabularx}

\usepackage{caption}

\usepackage{multirow}

\usepackage{multicol}

Lakshmaiah

\usepackage{graphicx}

\usepackage[document]{ragged2e}

\usepackage{xcolor}

\usepackage{pdfpages}

\usepackage{suffix}

\usepackage{draftwatermark}

Koneru

\SetWatermarkText{Koneru Lakshmaiah Educational Foundation}

\usepackage{fancyhdr}

\usepackage{pdflscape}

\renewcommand{\chaptername}{Experiment}

\oddsidemargin -0.1in

\textheight 9.2in

\textwidth 6.5in

\usepackage{fancyhdr}

\fancyhf{}

\fancyhead[L]{Research Methodology}

\fancyhead[R]{\leftmark}

\cfoot{\thepage}

Research Methodology EXPERIMENT 12. BOOK/THESIS-PREPARATION

\pagestyle{fancy}

\title{Research Methodology} %Compiled by:

\date{}

\author{IRD Team}

%Compiled by: \\ Dr. V Viswanath Shenoi \\ Professor, Department of Computer Science and

\newcommand\chapterauthor[1]{\authortoc{#1}\printchapterauthor{#1}}

\WithSuffix\newcommand\chapterauthor\*[1]{\printchapterauthor{#1}}

\makeatletter

%\newcommand{\nocontentsline}[3]{}

%\newcommand{\tocless}[2]{\bgroup\let\addcontentsline=\nocontentsline#1{#2}\egroup}

\newcommand{\printchapterauthor}[1]{%

{\parindent0pt\vspace\*{-25pt}%

\linespread{1.1}\large\scshape#1%

\par\nobreak\vspace\*{35pt}}

\@afterheading%

}

\newcommand{\authortoc}[1]{%

\addtocontents{toc}{\vskip-10pt}%

\addtocontents{toc}{%

Koneru Lakshmaiah Educational Foundation

\protect\contentsline{chapter}%

{\hskip1.3em\mdseries\scshape\protect\scriptsize#1}{}{}}

\addtocontents{toc}{\vskip5pt}%

}

\makeatother

\begin{document}

\include{RMRECORD/wrapper}

\maketitle

\tableofcontents

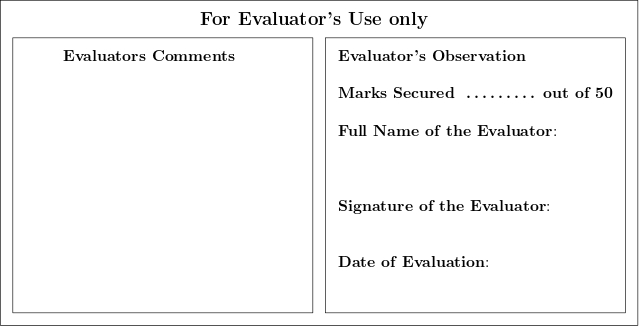
\include{RMRECORD/expt01}

\include{RMRECORD/expt02}

\include{RMRECORD/expt03}

\include{RMRECORD/expt04}

\end{document}

Research Methodology EXPERIMENT 12. BOOK/THESIS-PREPARATION

Koneru Lakshmaiah Educational Foundation

Research Methodology EXPERIMENT 12. BOOK/THESIS-PREPARATION

Koneru Lakshmaiah Educational Foundation

**Experiment 13**

**Ethical issues related to Publishing, Plagiarism and self-plagiarism**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

# DEFINITION OF PLAGIARISM

Plagiarism is presenting someone else’s work, including the work of other students, as one’s own. Any ideas or materials taken from another source for either written or oral use must be fully ac- knowledged, unless the information is common knowledge. What is considered “common knowl- edge”may differ from course to course.

Many people think of plagiarism as copying another’s work, or borrowing someone else’s original ideas. But terms like “copying”and “borrowing”can disguise the seriousness of the offense: Ac- cording to the Merriam-Webster OnLine Dictionary, to “plagiarize”means:

Koneru Lakshmaiah Educational Foundation

1. to steal and pass off (the ideas or words of another) as one’s own
2. to use (another’s production) without crediting the source
3. to commit literary theft
4. to present as new and original an idea or product derived from an existing source.

In other words, plagiarism is an act of fraud. It involves both stealing someone else’s work and lying about it afterward. But can words and ideas really be stolen? According to U.S. law, the answer is yes. In the United States and many other countries, the expression of original ideas is considered intellectual property, and is protected by copyright laws, just like original inven-

tions. Almost all forms of expression fall under copyright protection as long as they are recorded in some media.

###### All of the following are considered plagiarism:

* + turning in someone else’s work as your own.
  + copying words or ideas from someone else without giving credit
  + failing to put a quotation in quotation marks
  + giving incorrect information about the source of a quotation
  + changing words but copying the sentence structure of a source without giving credit

Research Methodology EXPERIMENT 13. ETHICAL ISSUES

* + - copying so many words or ideas from a source that it makes up the majority of your work, whether you give credit or not (see our section on “fair use” rules)

Attention! Changing the words of an original source is not sufficient to prevent plagiarism. If you have retained the essential idea of an original source, and have not cited it, then no matter how drastically you may have altered its context or presentation, you have still plagiarized Most cases of plagiarism can be avoided, however, by citing sources. Simply acknowledging that certain ma- terial has been borrowed, and providing your audience with the information necessary to find that source, is usually enough to prevent plagiarism.

# TYPES OF PLAGIARISM

* Word for Word Plagiarism: Copying a source’s text exactly the way it is written.
* Paraphrasing: Condensing the work of another in your owns words without citing.
* Copy & Paste Plagiarism: Copying and pasting text from an electronic source and using it as your own.
* Word Switch Plagiarism: Taking a sentence from a source and only switching around a few words.
* Style Plagiarism: Replicating the style and format of a source’s writing.
* Metaphor Plagiarism: Using metaphors or analogies from a source as your own.
* Idea Plagiarism: Using the creative ideas of another as your own.
* Self-Plagiarism: Reusing portions of previous writings in subsequent papers, either as a re- titled paper, or a compilation of bits and pieces of previous papers.

Koneru Lakshmaiah Educational Foundation

Anyone who has written or graded a paper knows that plagiarism is not always a black-and white issue. The boundary between plagiarism and research is often unclear. Learning to recognize the various forms of plagiarism, especially the more ambiguous ones, is an important step in the fight to prevent it.

##### SOURCES NOT CITED

1. “The Ghost Writer”The writer turns in another’s work, word-for-word, as his or her own.
2. “The Photocopy”The writer copies significant portions of text straight from a single source, without alteration.
3. “The Potluck Paper”The writer tries to disguise plagiarism by copying from several differ- ent sources, tweaking the sentences to make them fit together while retaining most of the original phrasing.
4. “The Poor Disguise”Although the writer has retained the essential content of the source, he or she has altered the paper’s appearance slightly by changing key words and phrases.
5. “The Labor of Laziness”The writer takes the time to paraphrase most of the paper from other sources and make it all fit together, instead of spending the same effort on original work.
6. “The Self-Stealer”The writer “borrows” generously from his or her previous work, violating policies concerning the expectation of originality adopted by most academic institutions.

Research Methodology EXPERIMENT 13. ETHICAL ISSUES

###### SOURCES CITED (but still plagiarized!)

1. “The Forgotten Footnote”The writer mentions an author’s name for a source, but neglects to include specific information on the location of the material referenced. This often masks other forms of plagiarism by obscuring source locations.
2. “The Misinformer”The writer provides inaccurate information regarding the sources, mak- ing it impossible to find them.
3. “The Too-Perfect Paraphrase”The writer properly cites a source, but neglects to put in quotation marks text that has been copied word-for-word, or close to it. Although attribut- ing the basic ideas to the source, the writer is falsely claiming original presentation and in- terpretation of the information.
4. “The Resourceful Citer”The writer properly cites all sources, paraphrasing and using quo- tations appropriately. The catch? The paper contains almost no original work! It is some- times difficult to spot this form of plagiarism because it looks like any other well-researched document.
5. “The Perfect Crime”Well, we all know it doesn’t exist. In this case, the writer properly quotes and cites sources in some places, but goes on to paraphrase other arguments from those sources without citation. This way, the writer tries to pass off the paraphrased mate- rial as his or her own analysis of the cited material.

# Preventing Plagiarism – Student Resources

In a research paper, you have to come up with your own original ideas while at the same time making reference to work that’s already been done by others. But how can you tell where their ideas end and your own begin? What’s the proper way to integrate sources in your paper? If you change some of what an author said, do you still have to cite that person? Confusion about the answers to these questions often leads to plagiarism. If you have similar questions, or are con- cerned about preventing plagiarism, we recommend using the checklist below.

Koneru Lakshmaiah Educational Foundation

1. Consult with your instructor Have questions about plagiarism? If you can’t find the an- swers on our site, or are unsure about something, you should ask your instructor. He or she will most likely be very happy to answer your questions. You can also check out the guide- lines for citing sources properly. If you follow them, and the rest of the advice on this page, you should have no problems with plagiarism.
2. Plan your paper Planning your paper well is the first and most important step you can take toward preventing plagiarism. If you know you are going to use other sources of informa- tion, you need to plan how you are going to include them in your paper. This means work- ing out a balance between the ideas you have taken from other sources and your own, orig- inal ideas. Writing an outline, or coming up with a thesis statement in which you clearly formulate an argument about the information you find, will help establish the boundaries between your ideas and those of your sources.
3. Take Effective Notes One of the best ways to prepare for a research paper is by taking thor- ough notes from all of your sources, so that you have much of the information organized before you begin writing. On the other hand, poor note-taking can lead to many problems – including improper citations and misquotations, both of which are forms of plagiarism!

Research Methodology EXPERIMENT 13. ETHICAL ISSUES

To avoid confusion about your sources, try using different colored fonts, pens, or pencils for each one, and make sure you clearly distinguish your own ideas from those you found elsewhere. Also, get in the habit of marking page numbers, and make sure that you record

bibliographic information or web addresses for every source right away – finding them again later when you are trying to finish your paper can be a nightmare!

1. When in doubt, cite sources Of course you want to get credit for your own ideas. And you don’t want your instructor to think that you got all of your information from somewhere else. But if it is unclear whether an idea in your paper really came from you, or whether you got it from somewhere else and just changed it a little, you should always cite your source. Instead of weakening your paper and making it seem like you have fewer original ideas, this will actually strengthen your paper by:
2. showing that you are not just copying other ideas but are processing and adding to them
3. lending outside support to the ideas that are completely yours, and
4. highlighting the originality of your ideas by making clear distinctions between them and ideas you have gotten elsewhere.
5. Make it clear who said what Even if you cite sources, ambiguity in your phrasing can of- ten disguise the real source of any given idea, causing inadvertent plagiarism. Make sure when you mix your own ideas with those of your sources that you always clearly distin- guish them. If you are discussing the ideas of more than one person, watch out for con- fusing pronouns. For example, imagine you are talking about Harold Bloom’s discussion of James Joyce’s opinion of Shakespeare, and you write: “He brilliantly portrayed the sit- uation of a writer in society at that time.”Who is the “He”in this sentence? Bloom, Joyce, or Shakespeare? Who is the “writer”: Joyce, Shakespeare, or one of their characters? Al-

Koneru Lakshmaiah Educational Foundation

ways make sure to distinguish who said what, and give credit to the right person. F Know’s how to Paraphrase: A paraphrase is a restatement in your own words of someone else’s ideas. Changing a few words of the original sentences does NOT make your writing a le- gitimate paraphrase. You must change both the words and the sentence structure of the original, without changing the content. Also, you should keep in mind that paraphrased passages still require citation because the ideas came from another source, even though you are putting them in your own words.

The purpose of paraphrasing is not to make it seem like you are drawing less directly from other sources or to reduce the number of quotations in your paper. It is a common miscon- ception among students that you need to hide the fact that you rely on other sources. Ac- tually it is advantageous to highlight the fact that other sources support your own ideas.

Using quality sources to support your ideas makes them seem stronger and more valid. Good paraphrasing makes the ideas of the original source fit smoothly into your paper, em- phasizing the most relevant points and leaving out unrelated information.

1. Evaluate Your Sources Not all sources on the web are worth citing – in fact, many of them are just plain wrong. So how do you tell the good ones apart? For starters, make sure you know the author(s) of the page, where they got their information, and when they wrote

it (getting this information is also an important step in avoiding plagiarism!). Then you should determine how credible you feel the source is: how well they support their ideas, the quality of the writing, the accuracy of the information provided, etc. We recommend using Portland Community College’s “rubrics for evaluating web pages” as an easy method of testing the credibility of your sources.

Research Methodology EXPERIMENT 13. ETHICAL ISSUES

# HOW TO AVOID PLAGIARISM

You use another person’s ideas, opinions, or theories.

* You use facts, statistics, graphics, drawings, music, etc., or any other type of information that does not comprise common knowledge.
* You use quotations from another person’s spoken or written word.
* You paraphrase another person’s spoken or written word.
* You draw from your own previous writing

## BEST PRACTICES TO AVOID PLAGIARISM

Writing Paraphrases or Summaries

* + - * Use a statement that credits the source somewhere in the paraphrase or summary
      * If you’re having trouble summarizing, try writing your paraphrase or summary of a text without looking at the original
      * Check your paraphrase or summary against the original text
      * Check your paraphrase or summary against sentence and paragraph structure
      * Put quotation marks around any unique words or phrases that you cannot or do not want to change Writing Direct Quotations
      * Keep the source’s name in the same sentence as the quote

Koneru Lakshmaiah Educational Foundation

* + - * Mark the quote with quotation marks
      * Quote no more material than is necessary
      * To shorten quotes by removing extra information, use ellipsis to indicate omitted text, but remember that: three ellipsis points indicates an in-sentence ellipsis, and four points for an ellipsis between two sentences
      * To add wording to a quote, place added words in brackets; be careful not to alter the origi- nal meaning of the quote
      * Use quotes that will have the most impact in your paper; too many direct quotes from sources may look like you have nothing to say

Writing About Another’s Ideas

* + - * Note the name of the idea’s originator in the sentence or throughout a paragraph about the idea
      * Use parenthetical citations, footnotes, or endnotes to refer readers to additional sources about the idea
      * Be sure to use quotation marks around phrases or words that the idea’s originator used to describe the idea

Revising, Proofreading, and Finalizing Your Paper

Research Methodology EXPERIMENT 13. ETHICAL ISSUES

* + - * + Proofread - check your notes and sources to make sure that anything coming from an out- side source is acknowledged in the following ways: o In-text citation o Footnotes or end- notes (if required) o Bibliography, References, or Works Cited page o Quotation marks around short quotes; longer quotes set off by themselves
        + If you have any questions about citation, ask your instructor BEFORE your paper is due to avoid having points deducted

THINGS TO AVOID WHEN WRITING

* + - * + Copying and pasting complete papers from electronic sources
        + Copying and pasting passages from electronic sources without placing the passages in quotes and properly citing the source
        + Having others write complete papers or portions of papers for you
        + Summarizing ideas without citing their source
        + Pulling out quotes from sources without putting quotation marks around the passages •

Closely paraphrasing

* + - * + Quoting statistics without naming the source
        + Using words and passages you don’t understand and can’t explain
        + Self-plagiarizing

Koneru Lakshmaiah Educational Foundation

* + - * + Making up sources
        + Making up citation information
        + Using photographs, video, or audio without permission or acknowledgment
        + Translating from one language to another without properly citing the original source
        + Not citing lectures, video chats, personal interviews, and other non-traditional source

**Self Plagiarism** This happens when the author has added research on a previously published article, book, contributed chapter, journal, and presents it as a new without acknowledging the fi rst article or taking permission from the previous publisher.Submission of the same article to multiple journals to increase the chances of publication or making multiple articles from a sin- gle article, known as, “salami slicing” is another form of plagiarism. WAME’s Ethics Committee says: “With respect to the issue of how much overlap is too much. . . a rule of thumb that some editors have applied when considering the amount of overlap between two review articles (not book chapters) has been overlap of more than one-third of the material.”

**Ghost writing** In this type the main contributor is not given due acknowledgement or some- one who has not contributed is given due credit. **Collusional** In this type the author asks a pro- fessional agent or institution to write an article and then claims as its own.

Research Methodology EXPERIMENT 13. ETHICAL ISSUES

# What is Citation?

A “citation” is the way you tell your readers that certain material in your work came from an- other source. It also gives your readers the information necessary to find that source again, in- cluding:

* information about the author
* the title of the work
* the name and location of the company that published your copy of the source
* the date your copy was published
* the page numbers of the material you are borrowing Why should I cite sources? Giving credit to the original author by citing sources is the only way to use other people’s work without plagiarizing. But there are a number of other reasons to cite sources:
  + Citations are extremely helpful to anyone who wants to find out more about your ideas and where they came from.
  + Not all sources are good or right – your own ideas may often be more accurate or in- teresting than those of your sources. Proper citation will keep you from taking the rap for someone else’s bad ideas.
  + Citing sources shows the amount of research you’ve done.
  + Citing sources strengthens your work by lending outside support to your ideas.

### Doesn’t citing sources make my work seem less original?

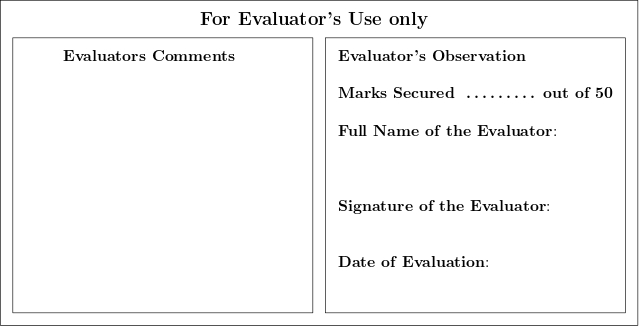
Koneru Lakshmaiah Educational Foundation

Not at all. On the contrary, citing sources actually helps your reader distinguish your ideas from those of your sources. This will actually emphasize the originality of your own work. When do I need to cite? Whenever you borrow words or ideas, you need to acknowledge their source. The following situations almost always require citation:

* + - * Whenever you use quotes
      * Whenever you paraphrase
      * Whenever you use an idea that someone else has already expressed
      * Whenever you make specific reference to the work of another
      * Whenever someone else’s work has been critical in developing your own ideas.

### How do I cite sources?

This depends on what type of work you are writing, how you are using the borrowed material, and the expectations of your instructor. First, you have to think about how you want to iden- tify your sources. If your sources are very important to your ideas, you should mention the au- thor and work in a sentence that introduces your citation. If, however, you are only citing the source to make a minor point, you may consider using parenthetical references, footnotes, or end- notes. There are also different forms of citation for different disciplines. For example, when you cite sources in a psychology paper you would probably use a different form of citation than you

Research Methodology EXPERIMENT 13. ETHICAL ISSUES

Koneru Lakshmaiah Educational Foundation

might in a paper for an English class. Finally, you should always consult your instructor to deter- mine the form of citation appropriate for your paper. You can save a lot of time and energy sim- ply by asking “How should I cite my sources,” or “What style of citation should I use?” before you begin writing. In the following sections, we will take you step-by-step through some general guidelines for citing sources.

**Experiment 14**

**Preparing Posters for project show**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

# 14.1 Introduction Poster

Foundation

Poster is a way of presenting our research ideas in the project show. In this exercise we shall learn on how the poster is prepared by using tikzposter class of overleaf. The following code will help to generate a poster with ease and consistency.

\documentclass[25pt, a0paper, portrait,margin=10mm, innermargin=10mm, blockverticalspace=7mm, colspace=10mm, subcolspace=5mm]{tikzposter}

Educational

\usepackage[utf8]{inputenc}

\makeatother

\title{\parbox{\linewidth}{\centering Title of the Project}}

\titlegraphic{\includegraphics{logo.png}}

% \date{\today}

Lakshmaiah

\usepackage{blindtext}

\usepackage{comment}

\usepackage{authblk}

\usepackage{graphicx}

\usepackage{color}

Koneru

\usepackage{xcolor}

\usepackage{verbatim}

\usetheme{Default}

\author[1]{Dr. G Murali}

\author[2]{Co-Principal Investigator 1}

\author[2]{Co-Principal Investigator 2}

\affil[1]{Afflitation 1}

\affil[2]{Affliation 2}

Research Methodology EXPERIMENT 14. POSTER

\makeatletter

\def\maketitle{\AB@maketitle}

\makeatother

\begin{document}

\maketitle

\begin{columns}

\column{0.5}

\block{Academic Year}{2024-25 }

\column{0.5}

\block{Project Funding reference No.}{ }

\end{columns}

\block{Name of the Funding Agency}{SERB

Foundation

}

\block{Graphical Abstract /Project Abstract / Introduction}{

}

\block{Framework (if any)}{

\begin{tikzfigure}

\includegraphics[width=0.4\textwidth]{logo.png}

\end{tikzfigure}

}

Educational

% \begin{columns}

% \column{0.5}

\block{Major Findings}{

\begin{itemize}

\item first point

\item second point

\end{itemize}

}

Lakshmaiah

% \column{0.5}

% \block{Limitations}{

% \begin{enumerate}

% \item first point

% \item second point

% \end{enumerate}

Koneru

% }

% \end{columns}

\block{Results}{

}

\block{Discussion \& Conclusion}{

}

\block{Publications from the project with Impact factor and category of journal}{

}

\block{References}{

}

\block{Acknowledgement}{

}

Research Methodology EXPERIMENT 14. POSTER

Koneru Lakshmaiah Educational Foundation

\end{document}

The output of the code is

Research Methodology EXPERIMENT 14. POSTER

Koneru Lakshmaiah Educational Foundation



Title of the project

Dr. G Murali1, Co-Principal Investigator 12, and Co-Principal Investigator 22 1Afflitation 1

2Affliation 2

**Academic Year**

**Project Funding reference No.**

2024-25

**Name of the Funding Agency**

SERB

**Graphical Abstract /Project Abstract / Introduction**

**Framework (if any)**

**Major Findings**

* first point
* second point

**Results**

**Discussion & Conclusion**

**Publications from the project with Impact factor and category of journal**

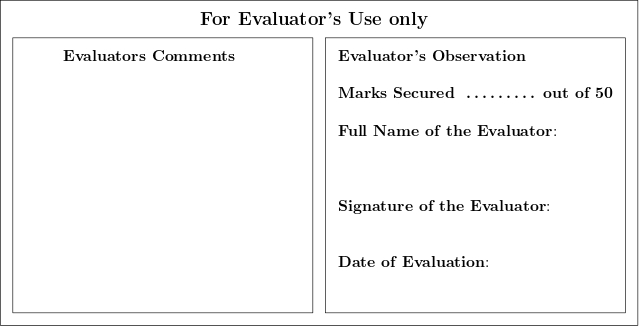
**References**

**Acknowledgement**

**LATEX** Ti*k* Z**poster**

Research Methodology EXPERIMENT 14. POSTER

Koneru Lakshmaiah Educational Foundation



Research Methodology EXPERIMENT 14. POSTER

Koneru Lakshmaiah Educational Foundation

**Experiment 15**

**Preparing Presentation in LATEX**

**Date of the Session**: *. . . . . . . . . . . . . . . . . .*

# Presentation

Presentations are a crucial tool in many fields, from academia to business, as they help commu- nicate ideas effectively to an audience. Here are some key points on the importance of presenta- tions:

* + 1. **Clarity of Communication**: Presentations allow complex ideas to be broken down and explained in a way that is easier to understand, using visuals, organized slides, and struc- tured flow.

Koneru Lakshmaiah Educational Foundation

* + 1. **Engagement**: A well-structured presentation can keep the audience’s attention, making the message more impactful. Elements like visuals, examples, and interactive discussions make content engaging.
    2. **Professionalism**: Presentations give a polished, professional format for sharing research, project updates, or business proposals, which enhances credibility and helps gain audience trust.
    3. **Persuasion and Influence**: Presentations are essential in fields where persuasion is key (like sales, marketing, and academia). A clear and confident presentation can influence de- cisions and convince others of the presenter’s viewpoint.
    4. **Knowledge Sharing**: Presentations facilitate the exchange of knowledge and ideas, allow- ing others to learn, provide feedback, and ask questions.
    5. **Developing Communication Skills**: Preparing and delivering presentations helps im- prove public speaking, organization, and the ability to explain complex information con- cisely, valuable skills for career growth.
    6. **Visual Impact**: Presentations leverage visual aids, which can enhance understanding, il- lustrate data, and emphasize key points, making the content memorable.

In summary, presentations are essential for effective communication, engagement, and influence, contributing to both personal and professional growth.

Research Methodology EXPERIMENT 15. SLIDES

# Creation of slides using LATEXBeamer class

To create the powerpoint, we need to to use beamer class. We need to chose any kind of theme, here we use AnnArbor theme.

% Creating a simple Title Page in Beamer

\documentclass{beamer}

% Theme choice:

\usetheme{AnnArbor}

% Title page details:

\title{Your First \LaTeX{} Presentation}

\subtitle{My-subtitle}

\author{latex-beamer.com}

\date{\today}

\begin{document}

% Title page frame

\begin{frame}

\titlepage

\end{frame}

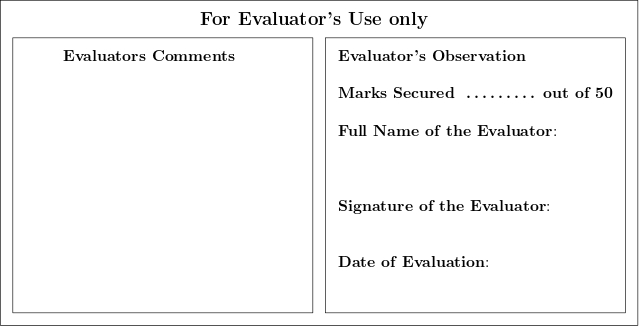
\end{document}

* + 1. We have chosen a predefined theme in Beamer, known as AnnArbor which is loaded using the command: \usetheme{AnnArbor}
    2. \title{}: is used to set a title to the presentation
    3. \subtitle{My-subtitle} to add the subtitle

Koneru Lakshmaiah Educational Foundation

* + 1. \author{}: is used to add authors’ names to the talk
    2. \date{}: is used to print the date of the talk, using \today will print the compilation day of the presentation.

The above code will generate: The slides are formed by means of \begin{frame}␣...␣\end{frame}

Research Methodology EXPERIMENT 15. SLIDES

Educational Foundation

Koneru Lakshmaiah

