

SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE
SUBJECT - Data Analysis with SAS / SPSS / R

Aim : Creating new variables using transformations and calculations in R. import dataset.

```
> library(dplyr)
> # 1. Import dataset
> titanic <- read.csv("titanic_submission.csv", na.strings = c("", "NA"))
> print("--- Original Titanic Data (First 6 Rows) ---")
[1] "--- Original Titanic Data (First 6 Rows) ---"
> print(head(titanic))
  PassengerId Survived
1         892         0
2         893         0
3         894         0
4         895         1
5         896         0
6         897         0
> # 2. Clean or handle missing values (e.g. Age or Fare)
> titanic_clean <- titanic %>%
+   mutate(
+     Age = ifelse(is.na(Age), median(Age, na.rm = TRUE), Age),
+     Fare = ifelse(is.na(Fare), median(Fare, na.rm = TRUE), Fare)
+   )
```