

Course Name: Statistics Using R Programming

Course Code: EMBSCBIO52

Duration: 1 hour daily- 1 month

To apply online : <http://www.eminentbio.com/home/Onlineform>

Trainer : Dr. Anuraj Nayariseri

<http://www.eminentbio.com/home/members>

https://www.researchgate.net/profile/Anuraj_Nayariseri

- **Descriptive Statistics**

- Mean, Median, Mode, Variance, Standard Deviation
- Range, Percentile, Quartiles, Inter Quartile Range
- Skewness and Kurtosis
- Central Moment

- **Basic Visualizations**

- Scatter plot, Boxplot, Barplot, Density plot

- **Inferential Statistical** (Parametric & Non-parametric tests)

- One Sample t-Test
- Wilcoxon Signed Rank Test
- Two Sample t-Test
- Wilcoxon Rank Sum Test
- Mann-Whitney-Wilcoxon Test
- Shapiro Test
- Kolmogorov And Smirnov Test
- Fisher's F-Test
- Chi Squared Test
- Correlation testing
- Fisher's exact test
- Annova
- **Outlier Treatment**
 - Detect outliers using univariate approach
 - Using multivariate approach
 - Cooks Distance
 - Outliers Test
 - Extract outliers using outliers package
 - Treating the outliers by Imputation/Capping/Prediction

- **Co-variance and Correlation**

- Pearson correlation
- Spearman correlation

- **Regression**

- Simple Linear regression
- Multiple Linear Regression
- Logistic regression
- Build Models
- Regression Diagnostics
- Predicting Models
- Prediction accuracy and error rates
- k- Fold Cross validation

- **Model fitness estimators**

- p Value: Checking for statistical significance
- test statistics
- R-Squared and Adj R-Squared
- Standard Error and F-Statistic
- AIC and BIC
- MSE (Mean squared error)
- MAPE (Mean absolute percentage error)

- **Multiple testing correction**

- **Multivariate data analysis**

- Principal Component Analysis (PCA)
- Multi Dimensional Scaling (MDS)
- Linear Discriminant analysis
- Correspondence analysis

- **Clustering**

- Hierarchical clustering
- k means clustering
- NMF clustering