



LESSON PLAN

LESSON PLAN INFORMATION

Subject/Course:	Data Science	Name:	Amit Arora
Grade:	High School	Date:	June 18 - July 18 2018
Topic:	R, Statistics, Linear Algebra, EDA, Visualizations, Machine Learning	Time:	90 - 120 minutes

Materials

- Slides for lesson
- Online resources

Lesson Plan

Introduction to R - I & II

- Setting up R Studio
- Data Types
- Loops and Conditional Statements
- Functions, vectorized version of functions
- Functional Programming
- Commonly used Packages and tasks: dplyr, purr, ggplot2, lubridate, tidyr, tidyverse
- Packages & Projects

Statistics & Linear Algebra (will be covered using R examples)

- The idea of distributions (continuous and discrete), why do we need models.
- Expected Value.
- Sampling, bootstrap and confidence intervals.
- The idea of statistical significance and associated tools and test like t test, Z-score, p-value etc
- Basic review of matrices, derivatives and related concepts to introduce the idea of optimization.
- Eigen values and Eigen vectors (time permitting).

Hunting for data, EDA and Visualization

- Using API provided by different websites to get data. We would be using AirNow API for the examples discussed in class.
- Exploratory Data Analysis using descriptive Statistics and Visualization.
- Explore use of Statistical distributions for modeling data.
- Study visualization using ggplot2 and plotly. Introduction to Grammar of Graphics.
- Use Maps, timeseries charts, density plots etc. for visualizations.

Machine Learning

- Use data sets discussed in earlier lessons to come up with prediction problems. Specifically understand (conceptually as well with examples)
- Linear Regression (use matrices and linear algebra from Lesson 2) for predicting continuous variable.
- Logistic regression for classification (time permitting understand from linear algebra perspective as well).
- K Means for clustering.

Potential datasets to be used for class examples

Air Quality Index

- See AirNow API: <https://docs.airnowapi.org/webservices>

BBC News

- See BBC News API: <https://newsapi.org/s/bbc-news-api>

Twitter

- See R package for Twitter API: <https://cran.r-project.org/web/packages/twitteR/twitteR.pdf>