# Course on Banking Analytics

Recently we were involved in creating and delivering a course on Banking Analytics for one of our customer.

The course was designed in such a way to have multiple channels for delivery. It was to be hosted on a learning management system so that interested people could take it on their own time and pace and also available as instructor led class room-based training.

This course is based on our Banking Analytics workshop offering. The purpose of this workshop is to to enrich the participants with the Analytics possibilities in addressing the challenges they face in their work. You can get more details of our offering at - <http://pexitics.com/analytics/resources/> and also contact us as [score@pexitics.com](mailto:score@pexitics.com) for further queries.

The engagement involved creating content on the banking processes and case studies around solving different problems using R/RStudio. The banking process’s around these banking area were briefly covered -

1. Commercial Banking Assets
2. Commercial Banking Liabilities
3. Non-Banking Financial Company (NBFC)
4. Insurance
5. Investment Banking
6. Risk
7. Brokerage
8. Intro to stats and R

The case studies in the above banking process was subdivided into these categories to cover and introduce different statistical methods –

1. Descriptive Statistics
2. Difference Statistics
3. Inferential Statistics
4. Associate Statistics
5. Predictive Statistics

All the case studies followed the “DCOVA and I” methodology.

# DCOVA and I

We used DCOVA & I framework to solve the case studies. In this framework analytics problem is broken down into these steps –

* **D**efine the Business Problem, the data that would be used and map the business problem into statistical problem
  + **C**ollect the data from appropriate sources as defined in the Define step
* Primary data
* Secondary Data
* Merge Different data sets
  + **O**rganize the data by classifying and combining structured and unstructured data into an analysable format
* Checking for and treating Missing Value
* Checking for and treating Outliers
* Creating Dummy Variables for factor variables
  + **V**isualize the data by developing different charts to find any trends in the data using
* Base R charts
* ggplot2
  + **A**nalyze the data by systematically applying statistical techniques.
  + **I**nsights – Insights by translating the outputs of Analyze steps to solve business problems

For solving business problems, just collecting data and running statistical operations on them is not sufficient. DCOVA with I is what will result in an accelerated positive change and strengthening of a competitive edge.

# Case Studies

The case studies were designed in such a way as to cover quite a few concepts including –

1. Description of the data
2. Checking and treatment for missing values
3. Checking and treatment for Outliers
4. Sampling and central limit theorem
5. Hypothesis Testing
6. Empirical and Chebyshev’s Theorem
7. Forecasting models
8. Probability of occurrence
9. Segmentation with clustering

# Sample case studies we covered

In this training, we solved over thirty case studies. All the case studies were solved using R/RStudio using DCOCA & I framework. I briefly describe some of the case studies here -

1. The new Risk manager wants to understand the portfolio data and find where the maximum values related to the portfolio lie. We used the concepts of normal distribution and checking for it by applying Shapiro-Wilk test, hypothesis testing and Chebyshev’s theorem.
2. A bank has tracked the cost of funds for various sources from 2003 to November 2016. The manager wants to understand if the cost of funds for all the sources are similar or are there differences in these. We used the concepts of hypothesis testing , t-test and anova in arriving at a conclusion.
3. We had a case study where the manager of the card issuing bank wanted to know the probability of a transaction being fraud given the amount band. Through this case study we introduced the concepts of conditional probability using Bayes theorem. In this case study, we used the concepts of probability and conditional probability.
4. In another case study, the manager wanted to Strengthen Fraud Detection Strategy during Credit Card Approval process. In this case study, we used historical data of the bank on customer information to build a model to identify fraudulent customers. We used Logistics Regression model to learn from historic data for prediction and then using the model to predict for a customer applying for a card.

# Shooting

For hosting the course into learning management system, we had to do video recording of the course.

The course was recorded at our customers studios. Most of the course was recorded by Subhashini and Reuben (the cofounders of Pexitics ).

I was involved in recording for one case study which was missed out initially. I found this activity to be the most challenging part in the whole engagement. For me, teaching in an empty room facing a monitor was not easy.

After this experience, I appreciate more the instructors on the MOOC courses, my only experience in taking courses on LMS !