Dr Milan Joshi

# 1. Bank Revenues

**MULTIPLE REGRESSION - SEVERAL PREDICTORS, WITH TRANSFORMATION**

**KEY IDEAS:**
The log transformation, stepwise regression, regression assumptions, residuals, Cook's D, interpreting model coefficients, singularity, Prediction Profiler, inverse transformations.

## BACKGROUND:
A bank wants to understand how customer banking habits contribute to revenues and profitability.

## THE TASK:
We want to build a model that allows the bank to predict profitability for a given customer. The resulting model will be used to forecast bank revenues and guide the bank in future marketing campaigns.

## 2. Titanic Passengers

**LOGISTIC REGRESSION**

**KEY IDEAS:**
Logistic regression, log odds and logit, odds, odds ratios, prediction profiler.

**BACKGROUND:**
The sinking of the RMS Titanic is one of the most infamous shipwrecks in history. On April 15, 1912, during her maiden voyage, the Titanic sank after colliding with an iceberg, killing 1502 out of 2224 passengers and crew.

**THE TASK:**
We use this rich and storied example to explore some questions of interest about survival rates for the Titanic. For example, were there some key characteristics of the survivors? Were some passenger groups more likely to survive than others? Can we accurately predict survival?

## 3. Credit Card Marketing

**CLASSIFICATION TREE**
**KEY IDEAS:**
Classification trees, validation, confusion matrix, misclassification, leaf report, ROC curves, lift curves.

**BACKGROUND:**
A bank would like to understand the demographics and other characteristics associated with whether a customer accepts a credit card offer.

**THE TASK:**
We want to build a model that will provide insight into why some bank customers accept credit card offers. Because the response is categorical (either Yes or No) and we have a large number of potential predictor variables, we use the Partition platform to build a classification tree for Offer Accepted.

**4.** Customer Churn

**NEURAL NETWORKS IN JMP PRO**

**KEY IDEAS:**

Neural networks, activation functions, model validation, confusion matrix, lift, prediction profiler, variable importance.

**BACKGROUND:**

Customer retention is a challenge in the ultracompetitive mobile phone industry. A mobile phone company is studying factors related to customer churn, a term used for customers who have moved to another service provider.

**THE TASK:**

The company would like to build a model to predict which customers are most likely to move their service to a competitor. This knowledge will be used to identify customers for targeted interventions, with the ultimate goal of reducing churn.

**5.**Boston Housing

**MODEL VALIDATION AND COMPARISON IN JMP PRO**

**KEY IDEAS:**

Model validation, stepwise regression, regression trees, neural networks, validation statistics and model comparison.

**BACKGROUND:**

The objective of this study is to develop a model to predict the median value of homes in the Boston area.

**THE TASK:**

Our goal is to use the available data build a model that makes accurate predictions about home values in the Boston area. To ensure that the model predicts well for data not used to build the model, we use model validation. We will build different models (e.g., multiple regression, regression tree and neural network), compare the performance of these models, and select the best-performing model.