Prediction of Bike Rental using Model Reuse Strategy

Arun Bala Subramaniyan
Dr. Rong Pan
Arizona State University, USA

About Myself

- M.S. Industrial Engineering (2014 - 2016)

  - School of Computing, Informatics, Decision Systems Engineering, ASU

- Nanodegree in Data Science (2015 - 2016)
- Research:

  - Reliability Based Robust Design Optimization (RBRDO)

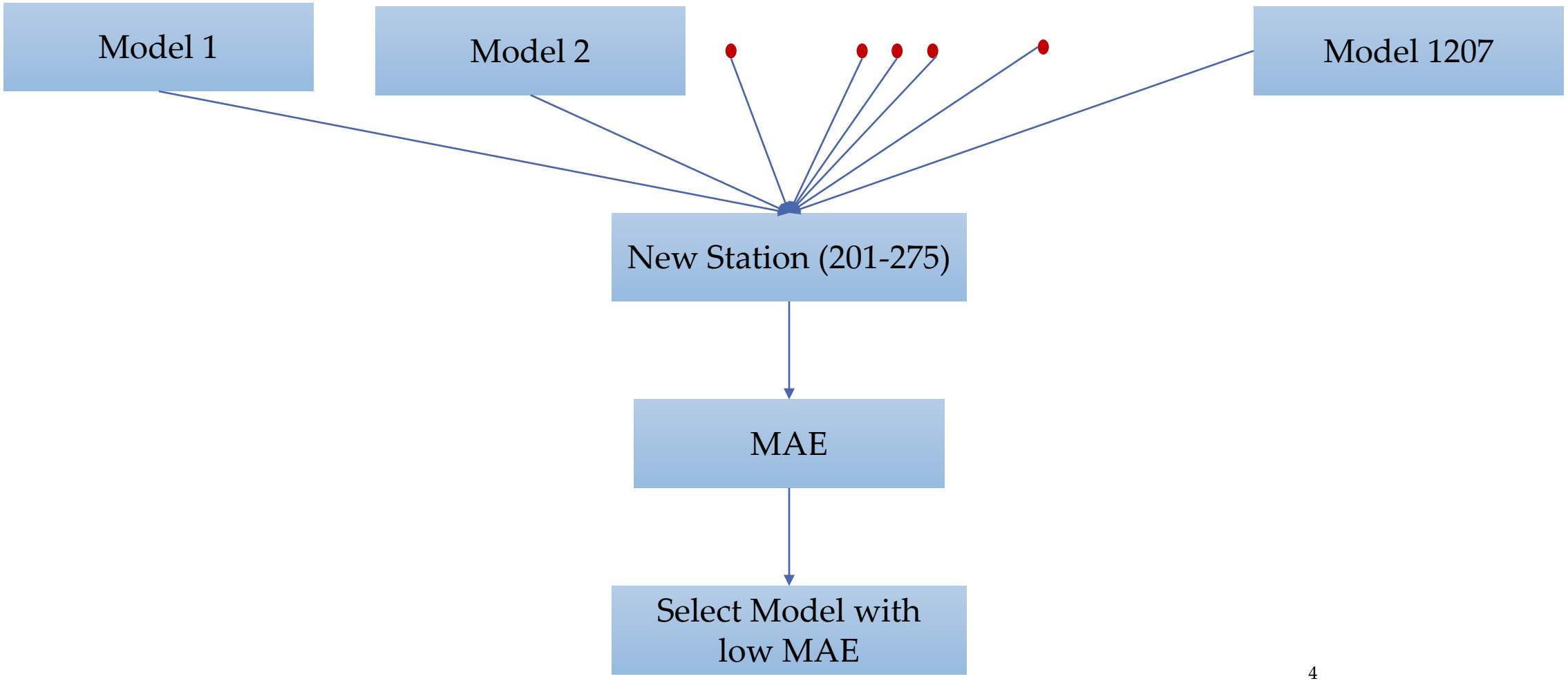
- Future interests:

  - Ph.D.
  - Entrepreneurship

MoReBikeS 2015

- Objective
  - To predict the number of bikes in new stations
- Methodology
  - Reuse the models learnt from old stations (1 to 200)
- Extraction algorithm
  - 7 base models+(6*200 trained models) = 1207 models

Model Extraction



MoReBikeS 2015

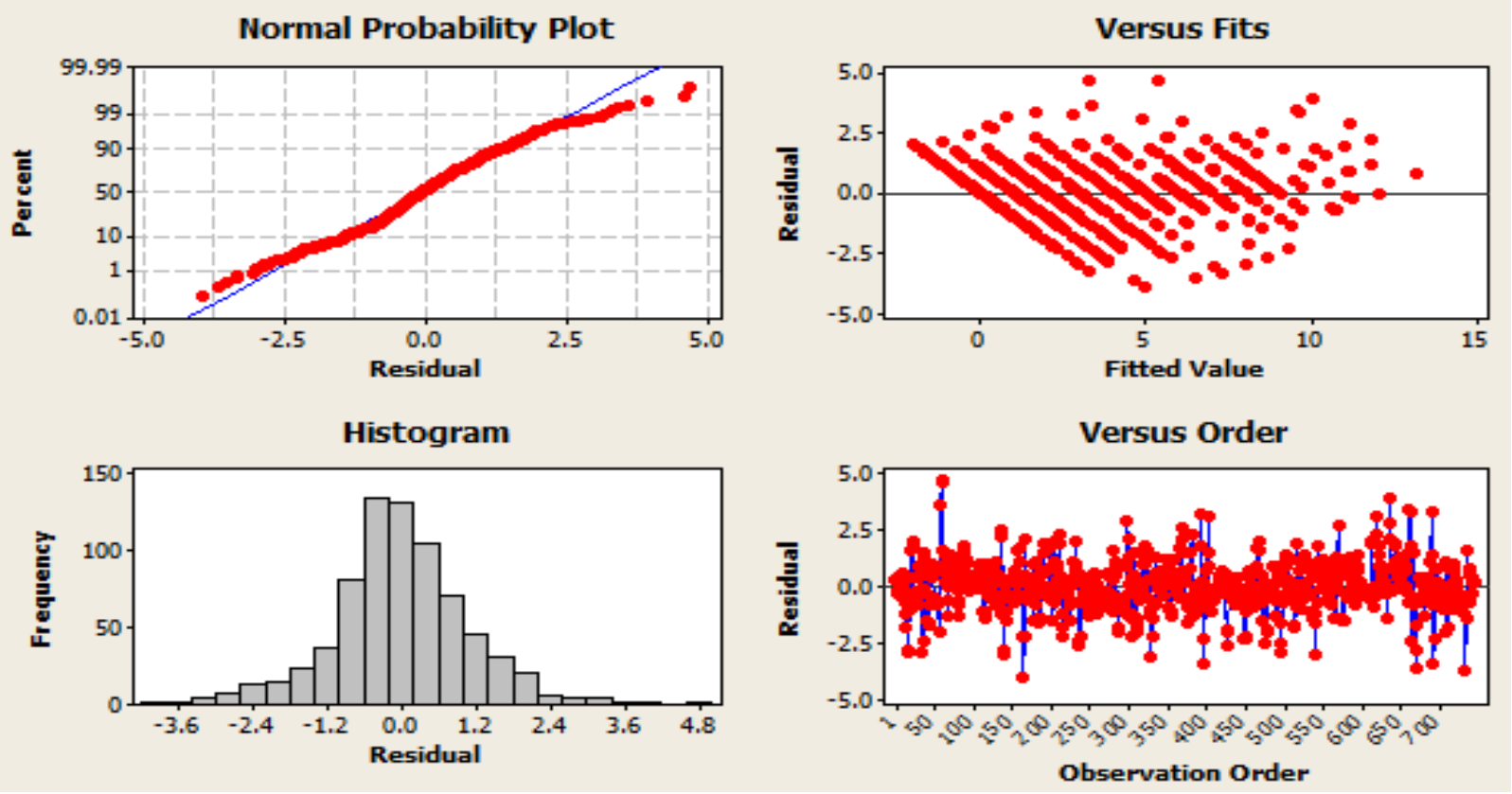
- Objective
  - To predict the number of bikes in new stations
- Methodology
  - Reuse the models learnt from old stations (1 to 200)
- Extraction algorithm
  - 7 base models+(6*200 trained models) = 1207 models
  - Constraint: [0,max_docks]
- Prediction algorithm
- MAE = 2.502 (Small test set leaderboard)
- MAE = 2.067 (Full test set leaderboard)

Methods tried

- Ordinary Least Squares (OLS) Method
- Poisson Regression
- Zero Inflated Poisson (ZIP) Regression

Ordinary Least Squares (OLS)

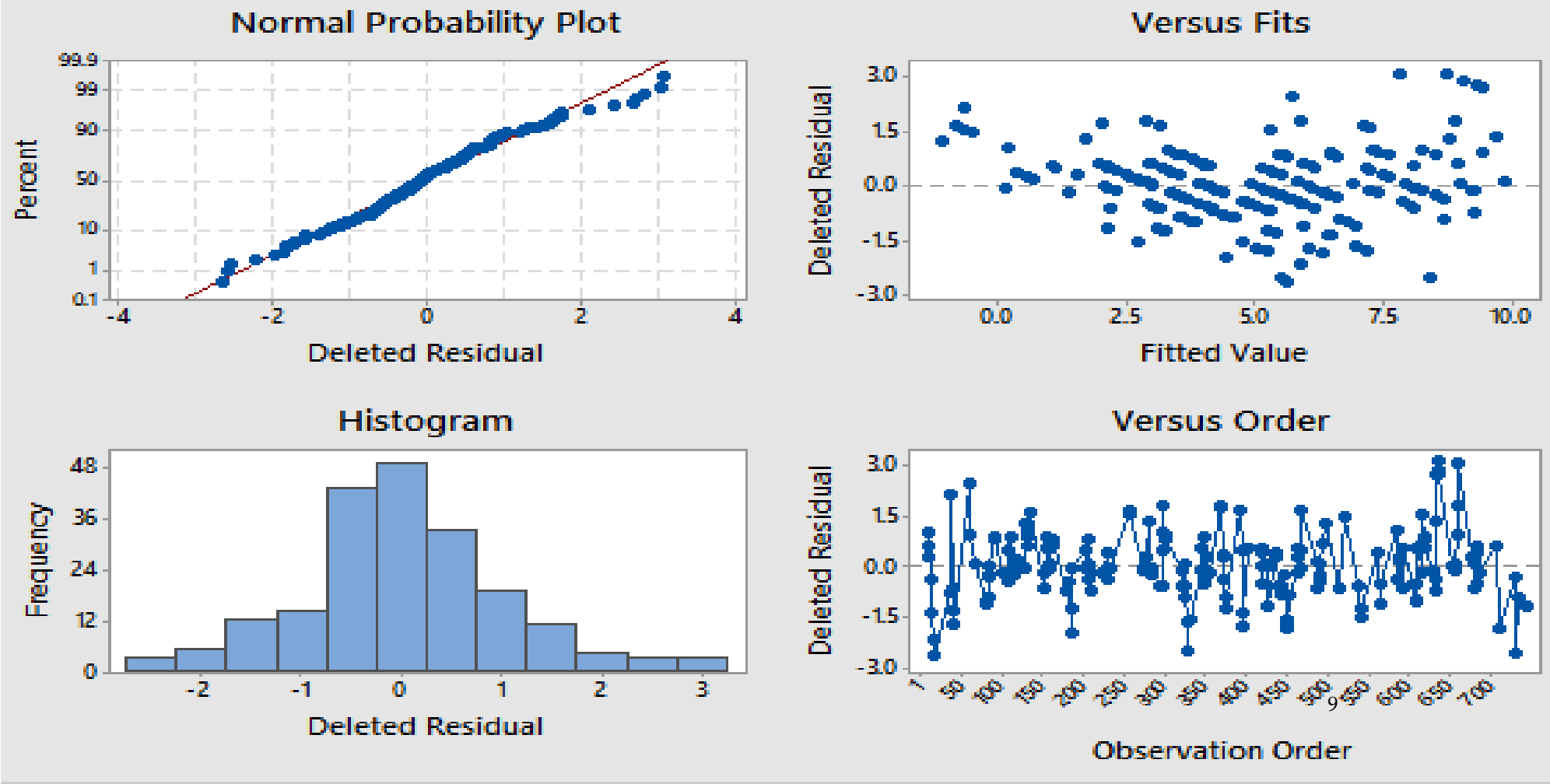
- Initial model



Ordinary Least Squares (OLS)

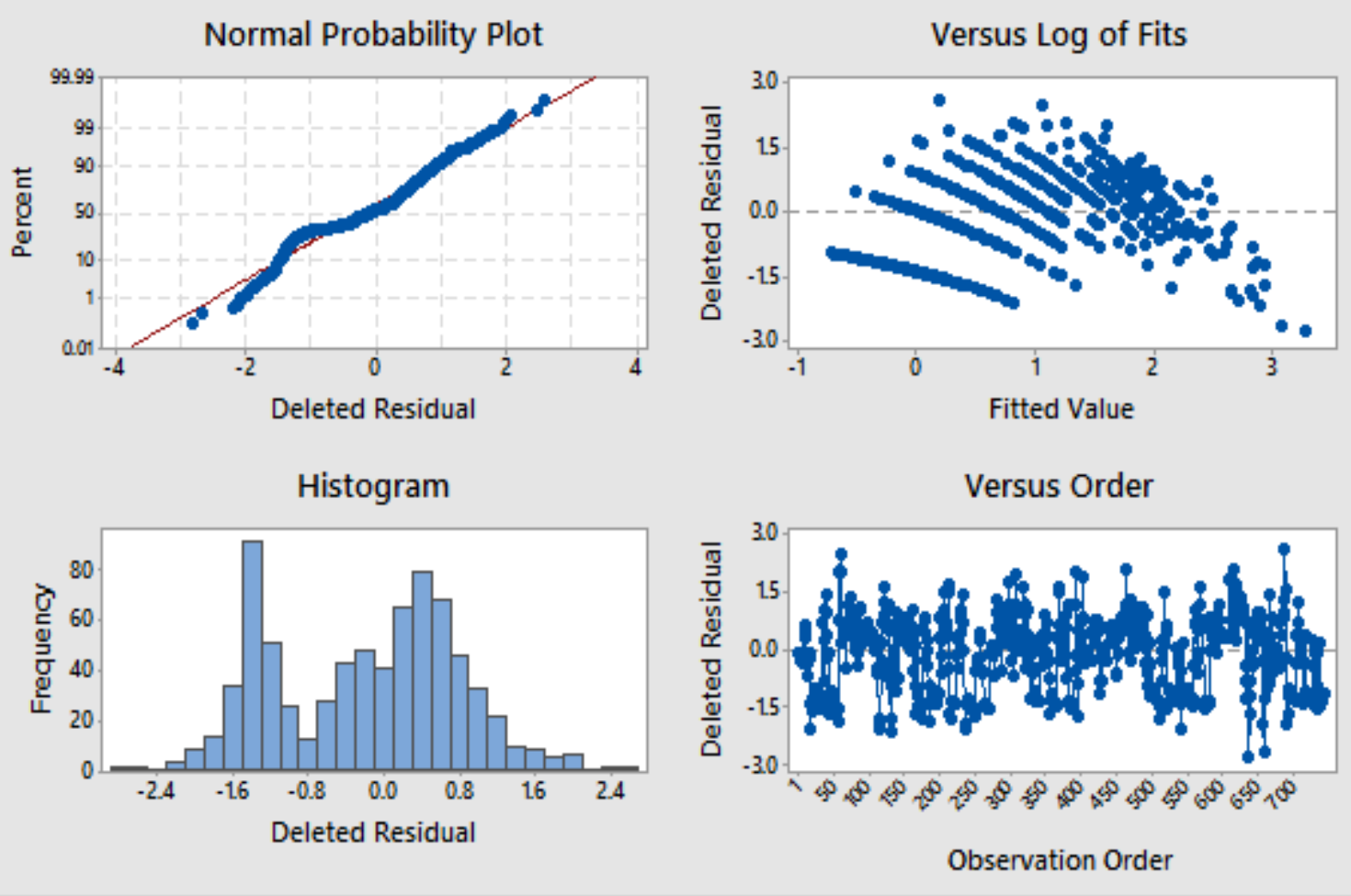
- Initial model
- Subset Regression
- Forward and Backward Elimination
- Transformation and polynomial regression
- Final Model
- MAE: 2.724

Final Model - OLS



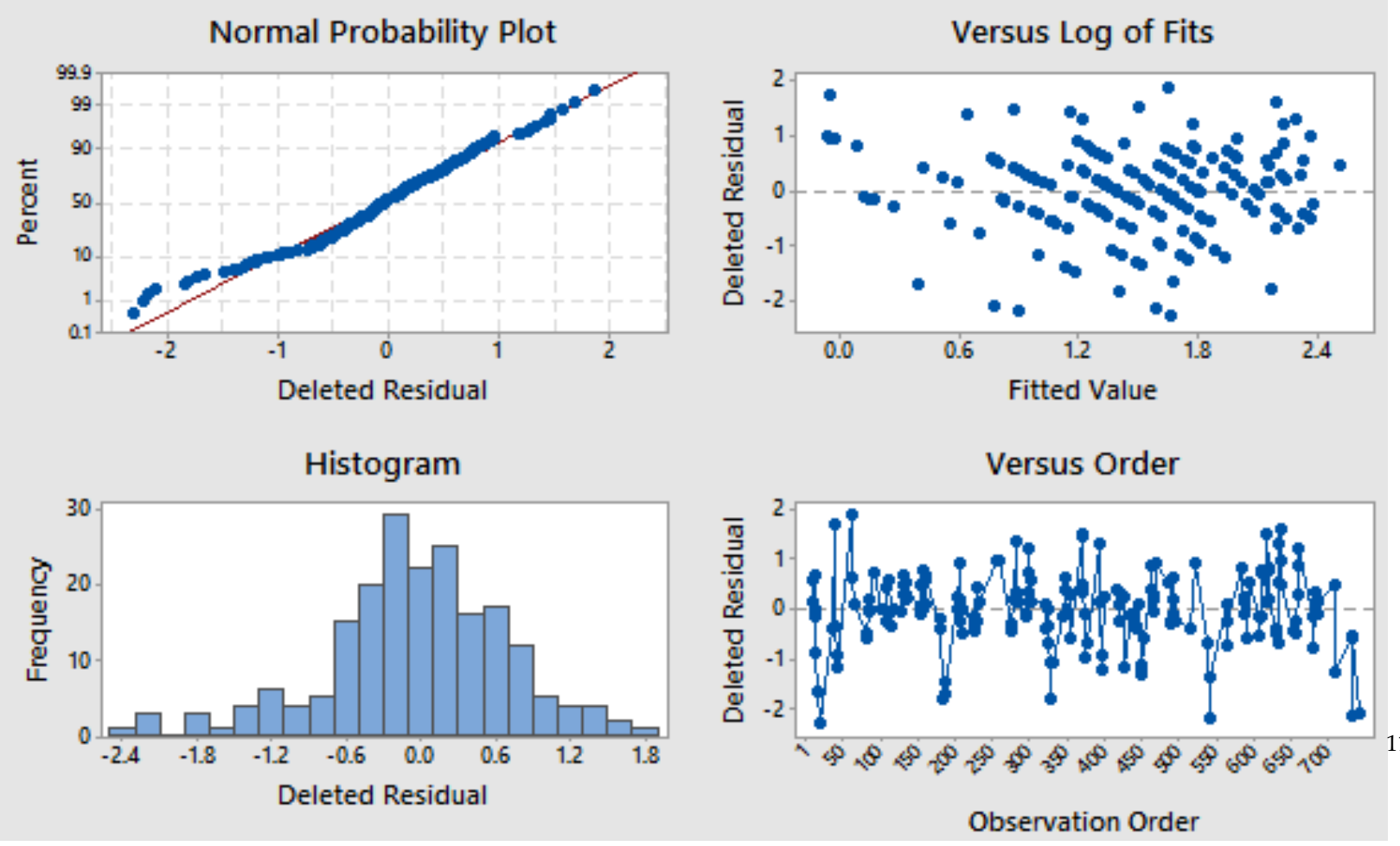
Poisson Regression

- Initial Model



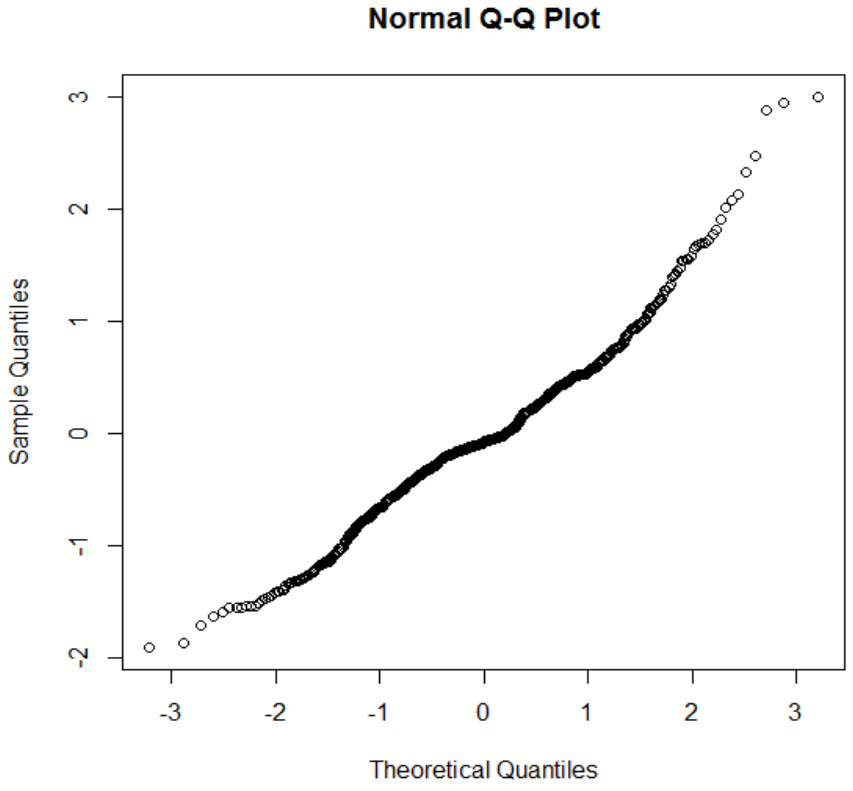
Final Model – Poisson Regression

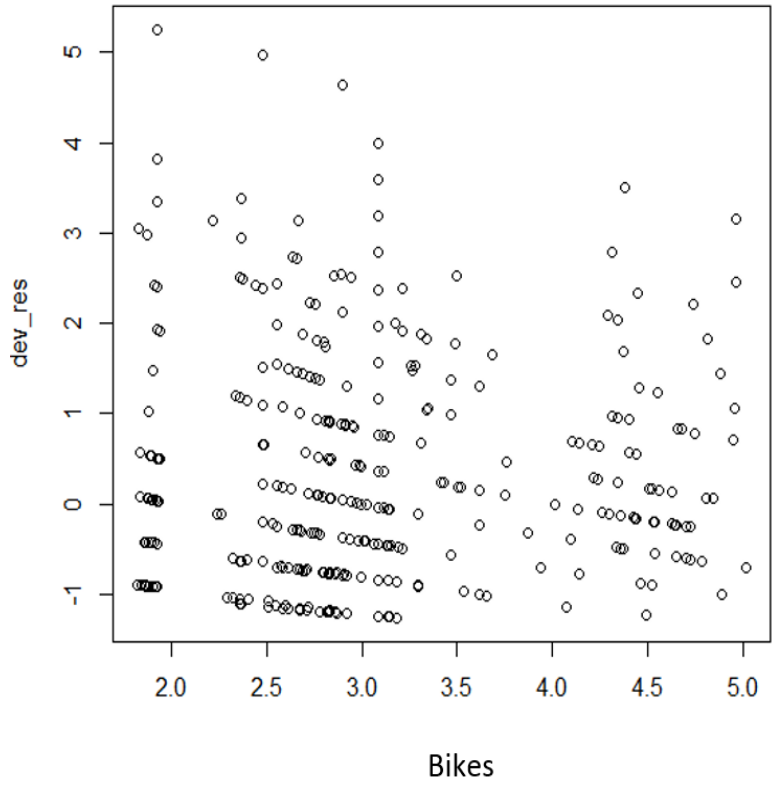
MAE: 3.068



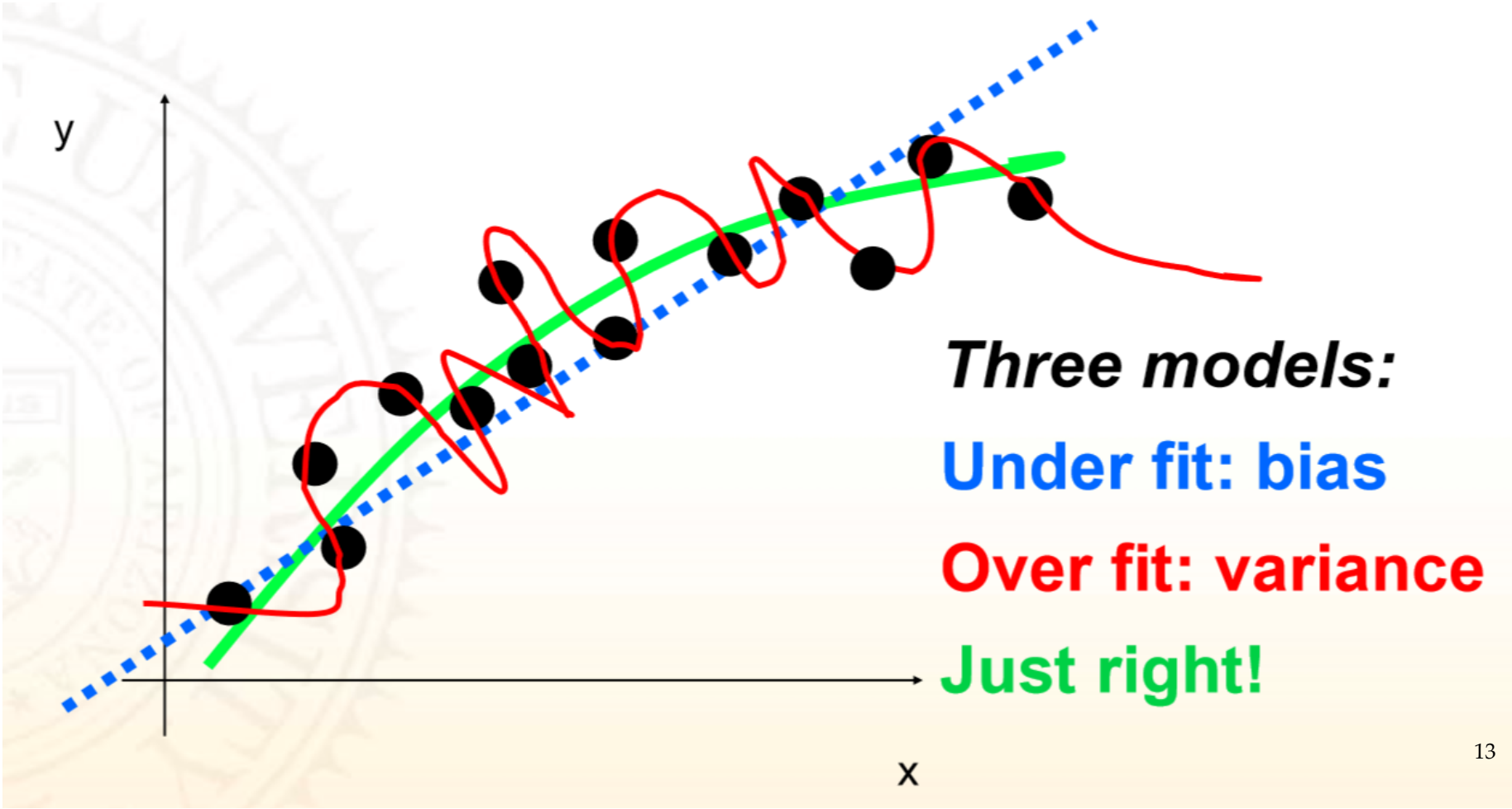
Zero Inflated Poisson Regression

MAE: 2.774





My Favorite Chart!!
-Bias (vs) Variance



Recommendation

- LASSO Regression
- Location Based Sampling
- Modelling the error values

Questions??