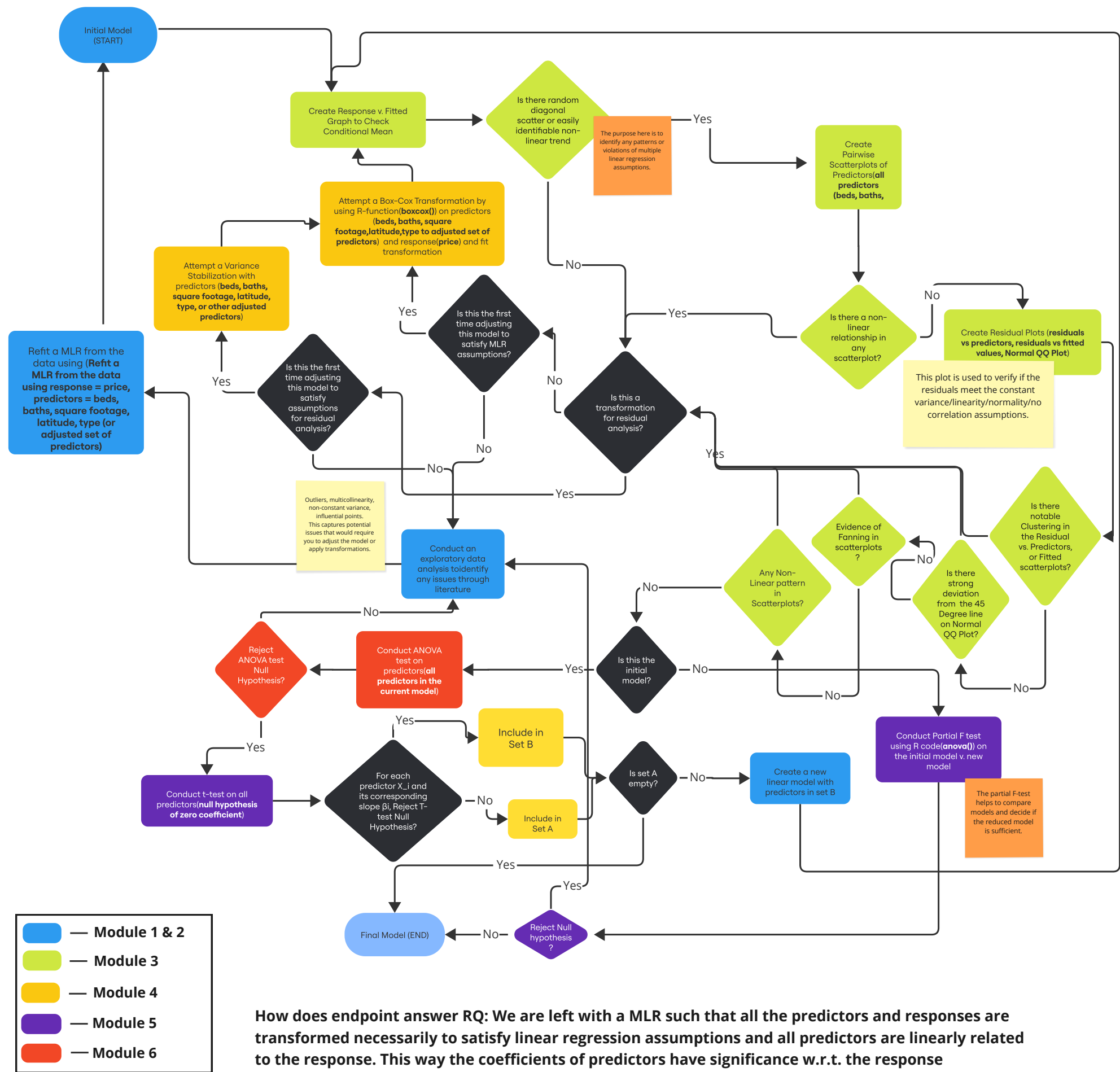**Investigation into the factors influencing Toronto's rental prices over the decades, using data from property listings to explore how aspects like property type, size, and location impact rents.**

Dai - General Framework for Flowchart
Bu- Editing and organizing Flowchart
Uma - Colouring and Editing Flowchart

Initial Model (START)

Create Response v. Fitted Graph to Check Conditional Mean

Is there random diagonal scatter or easily identifiable non-linear trend

The purpose here is to identify any patterns or violations of multiple linear regression assumptions.

Create Pairwise Scatterplots of Predictors(**all predictors (beds, baths,**

Attempt a Box-Cox Transformation by using R-function(**boxcox()**) on predictors (**beds, baths, square footage,latitude,type to adjusted set of predictors**) and response(**price**) and fit transformation

Attempt a Variance Stabilization with predictors (**beds, baths, square footage, latitude, type, or other adjusted predictors**)

Is this the first time adjusting this model to satisfy MLR assumptions?

Is there a non-linear relationship in any scatterplot?

Create Residual Plots (**residuals vs predictors, residuals vs fitted values, Normal QQ Plot**)

This plot is used to verify if the residuals meet the constant variance/linearity/normality/no correlation assumptions.

Refit a MLR from the data using (**Refit a MLR from the data using response = price, predictors = beds, baths, square footage, latitude, type (or adjusted set of predictors)**)

Is this the first time adjusting this model to satisfy assumptions for residual analysis?

Is this a transformation for residual analysis?

Is there notable Clustering in the Residual vs. Predictors, or Fitted scatterplots?

Outliers, multicollinearity, non-constant variance, influential points. This captures potential issues that would require you to adjust the model or apply transformations.

Conduct an exploratory data analysis to identify any issues through literature

Evidence of Fanning in scatterplots?

Any Non-Linear pattern in Scatterplots?

Is there strong deviation from the 45 Degree line on Normal QQ Plot?

Is this the initial model?

Reject ANOVA test Null Hypothesis?

Conduct ANOVA test on predictors(**all predictors in the current model**)

Include in Set B

Is set A empty?

Create a new linear model with predictors in set B

Conduct Partial F test using R code(**anova()**) on the initial model v. new model

Conduct t-test on all predictors(**null hypothesis of zero coefficient**)

For each predictor X_i and its corresponding slope βi, Reject T-test Null Hypothesis?

Include in Set A

The partial F-test helps to compare models and decide if the reduced model is sufficient.

Final Model (END)

Reject Null hypothesis?

Legend:
— Module 1 & 2
— Module 3
— Module 4
— Module 5
— Module 6

How does endpoint answer RQ: We are left with a MLR such that all the predictors and responses are transformed necessarily to satisfy linear regression assumptions and all predictors are linearly related to the response. This way the coefficients of predictors have significance w.r.t. the response

Yes · No