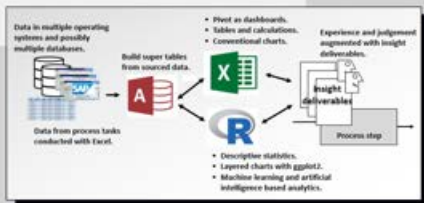# Data and Analytics Skills
for
Your Career Security

Keeping it simple. . .
only the skills you're likely to use

Richard G. Lamb

Images in color

Note: Book is produced in grayscale but color is part of the story for some graphic outputs. Such outputs are contained in the pdf.
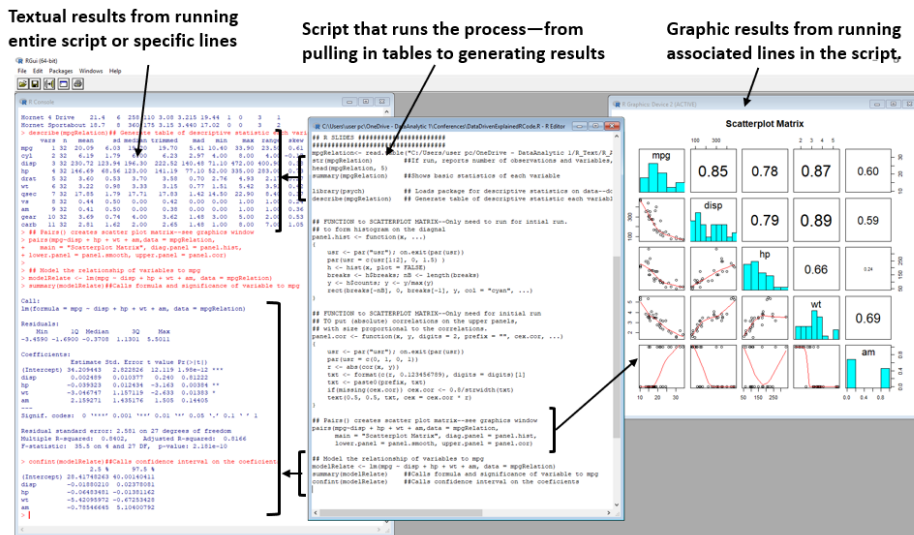
Textual results from running entire script or specific lines

Script that runs the process—from pulling in tables to generating results

Graphic results from running associated lines in the script.

Figure 2-9: Console, script and graphic windows of R.

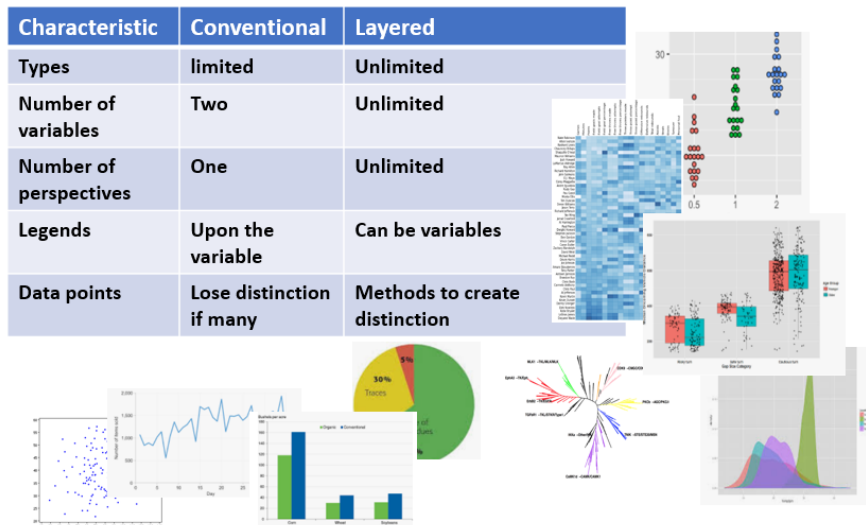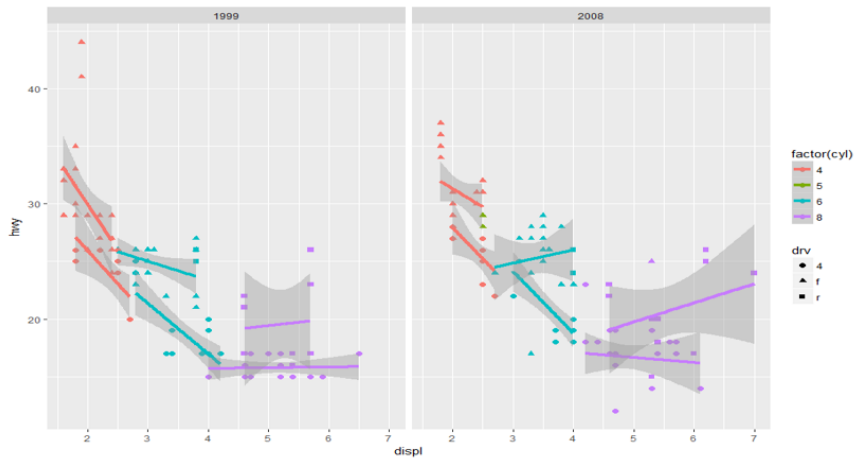| Characteristic	Conventional	Layered
Types	limited	Unlimited
Number of variables	Two	Unlimited
Number of perspectives	One	Unlimited
Legends	Upon the variable	Can be variables
Data points	Lose distinction if many	Methods to create distinction

Figure 2-10: Contrasts between traditional and layered charting.

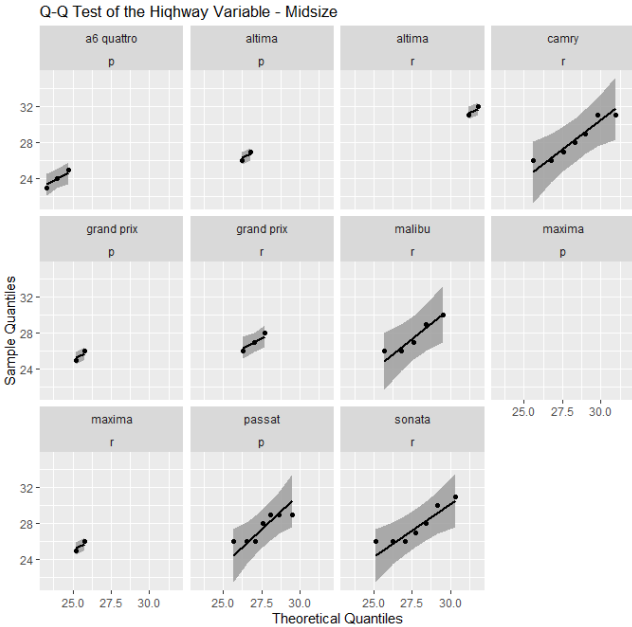**Figure 2-11: Layered charting to present measures of performance.**
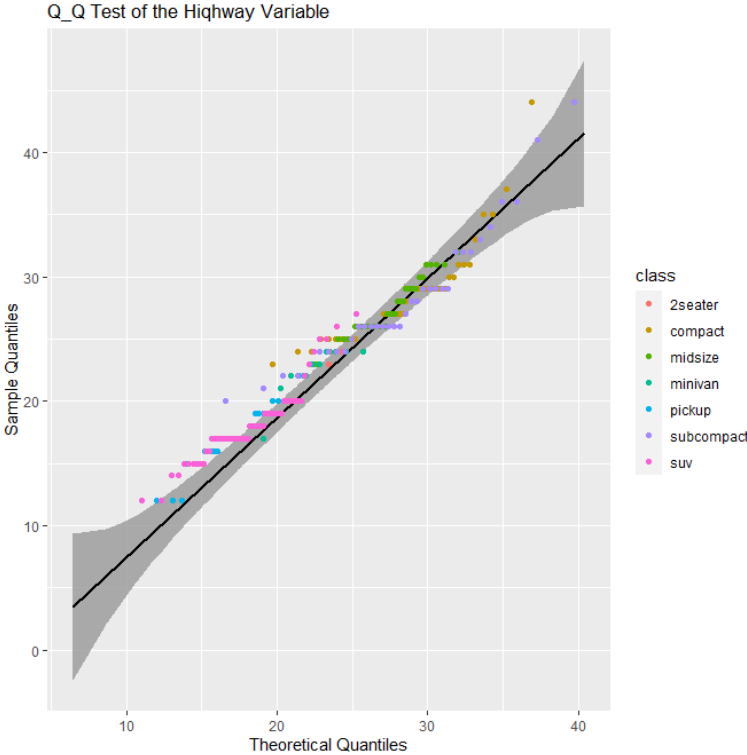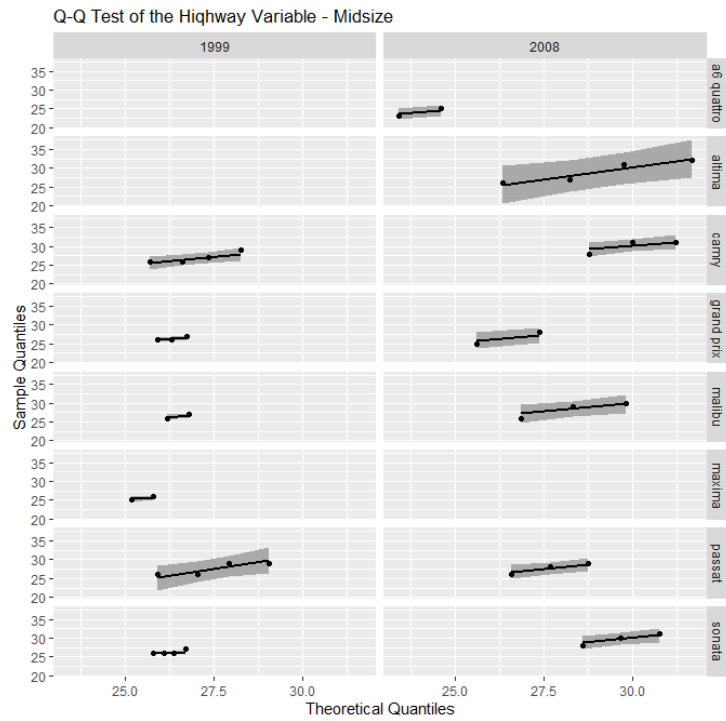
Output 5-9: Q-Q plot of the highway mileage variable.
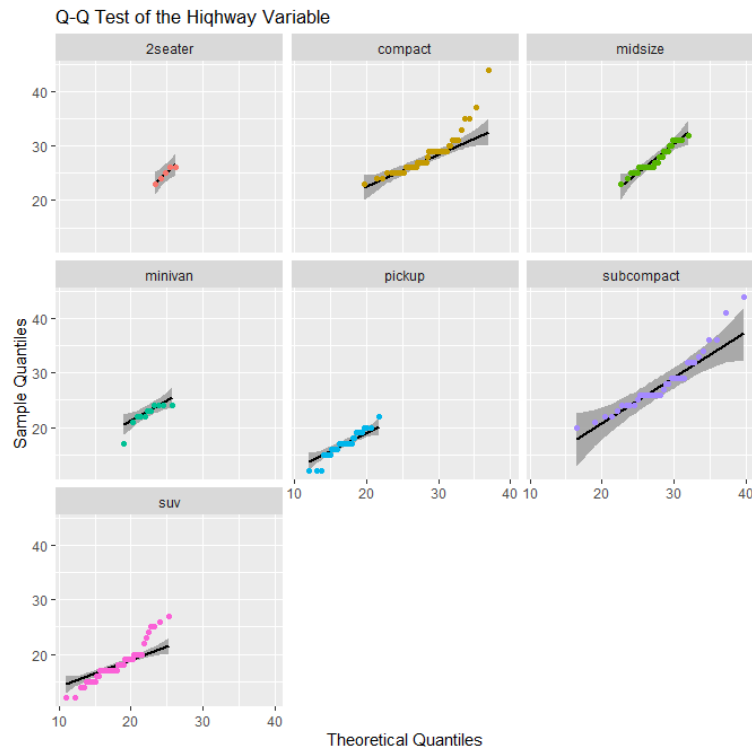
Output 5-11: Distribution of the hwy variable subset upon class.

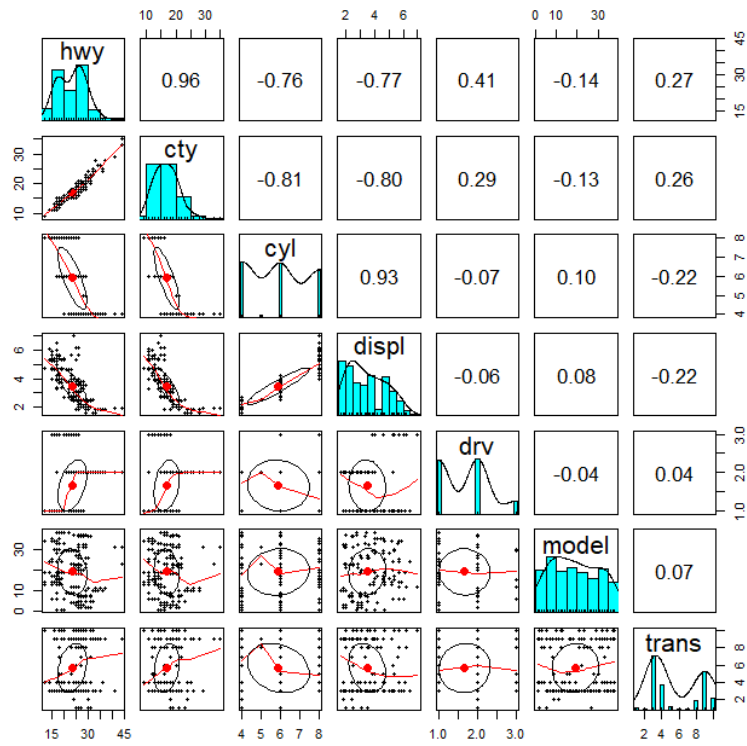Q-Q Test of the Hiqhway Variable - Midsize
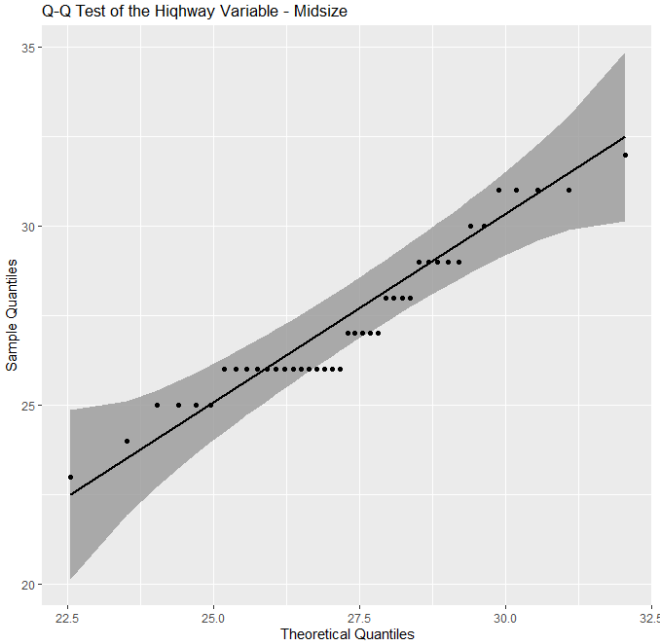
Output 5-12: The midsize class does not show normal distribution.
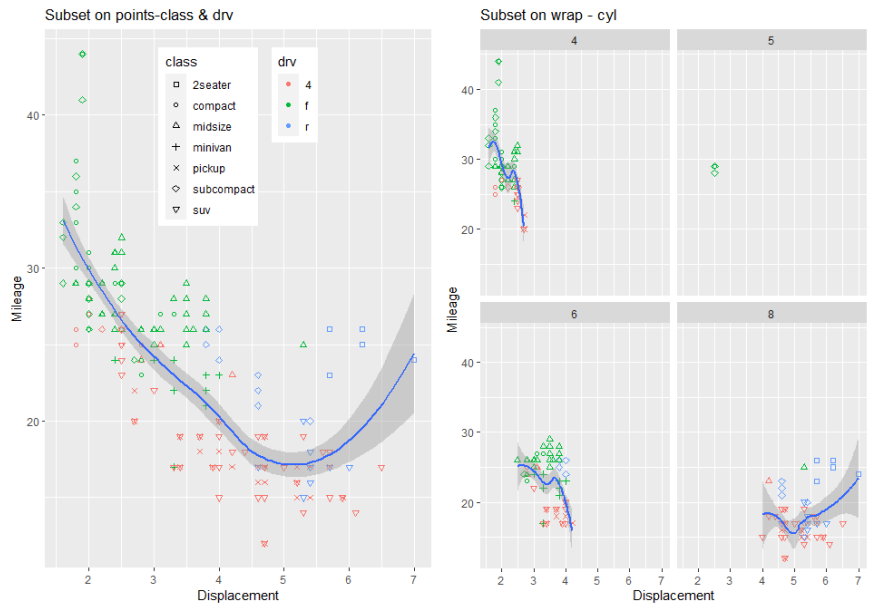
Output 5-13: Subsetted facets on model and fl to the Q-Q test of the highway variable.
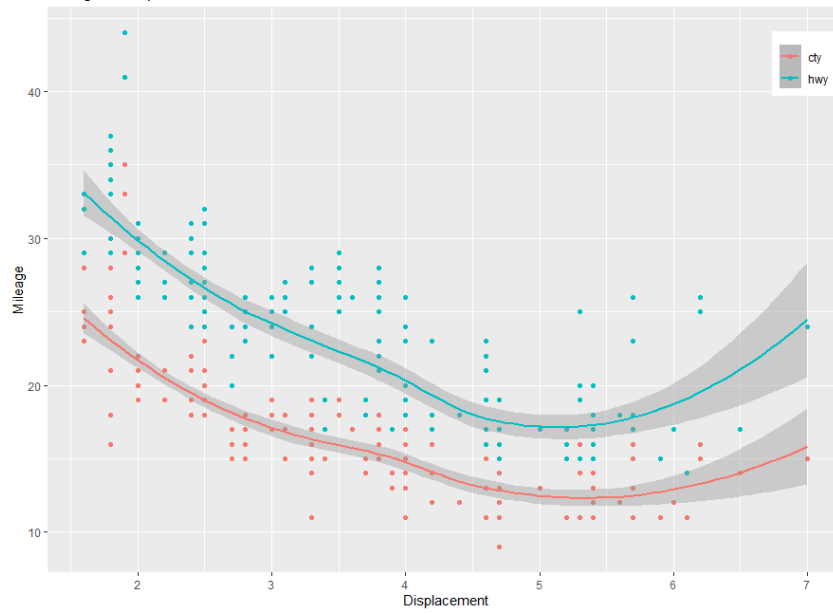
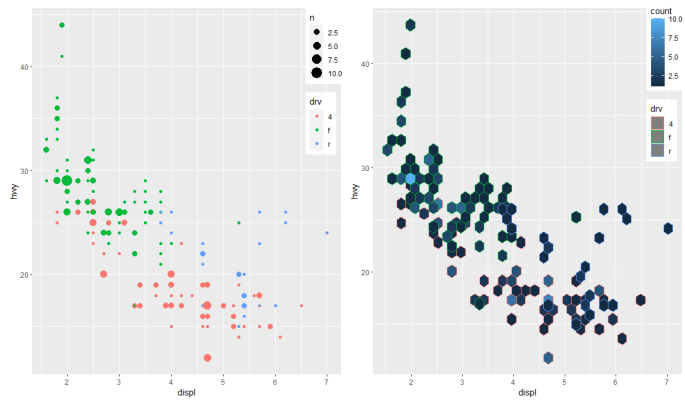**Output 5-14: Q-Q test by model and year.**

Output 5-15: Grid of correlations and more to mpg variables.
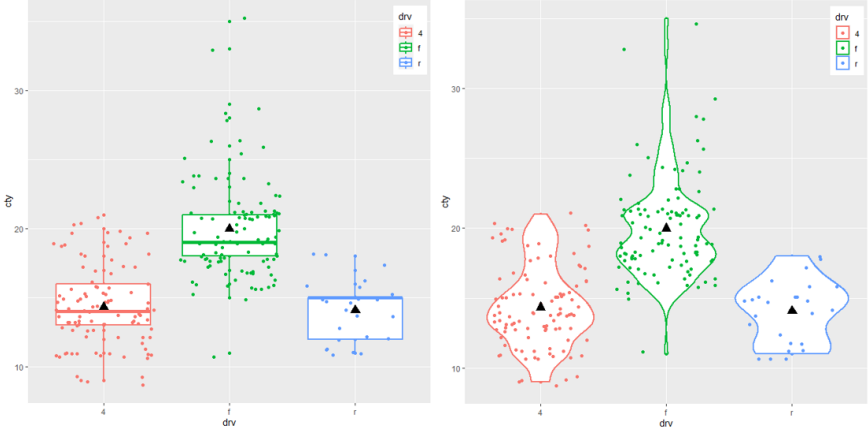
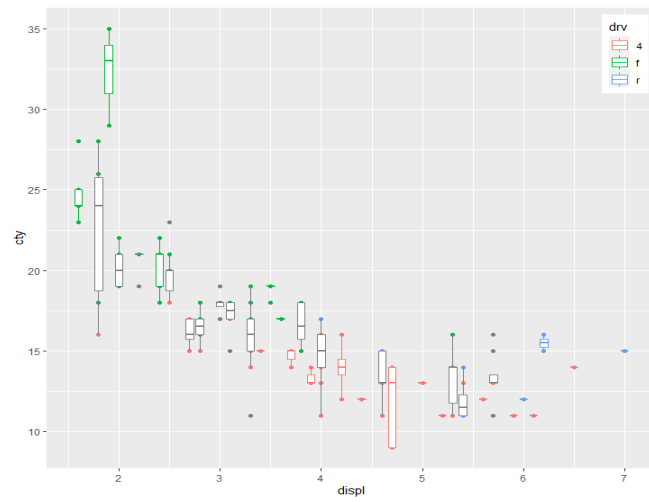**Output 5-17: Scatter plots subset on points and facets.**

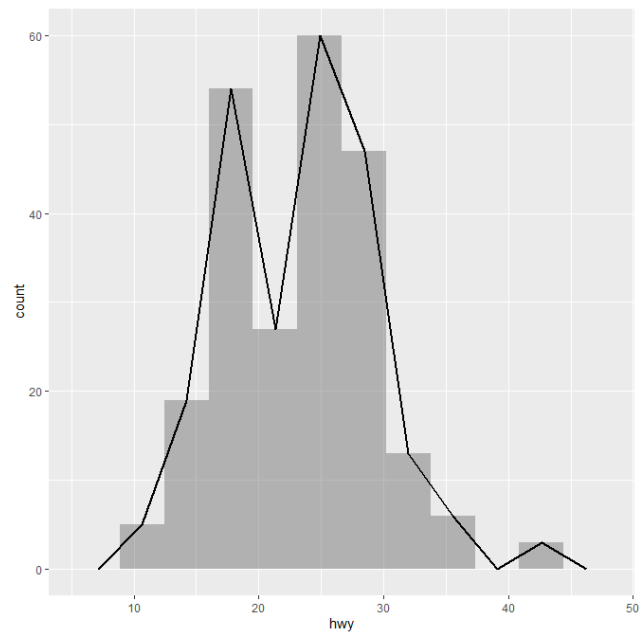**Output 5-18: Two scatter graphs presented as a single graph.**

Output 5-19: Area count and hex methods to deal with over plotting points.
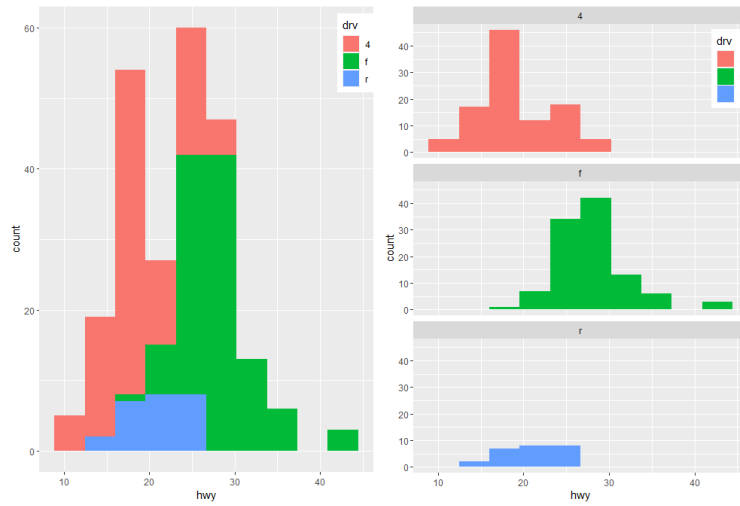
Output 5-20: Box and violin plots to compare central tendency and , variance among data, and spot outliers.
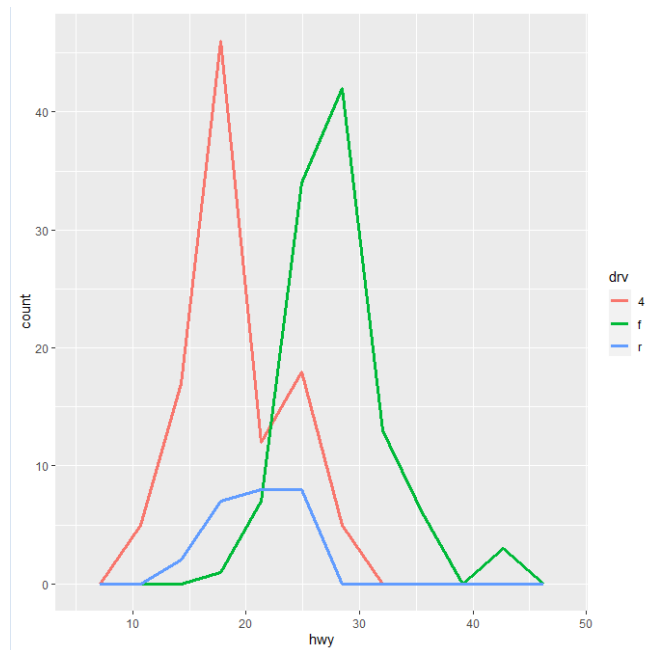
Output 5-21: Boxplot with a continuous variable as its grouping.
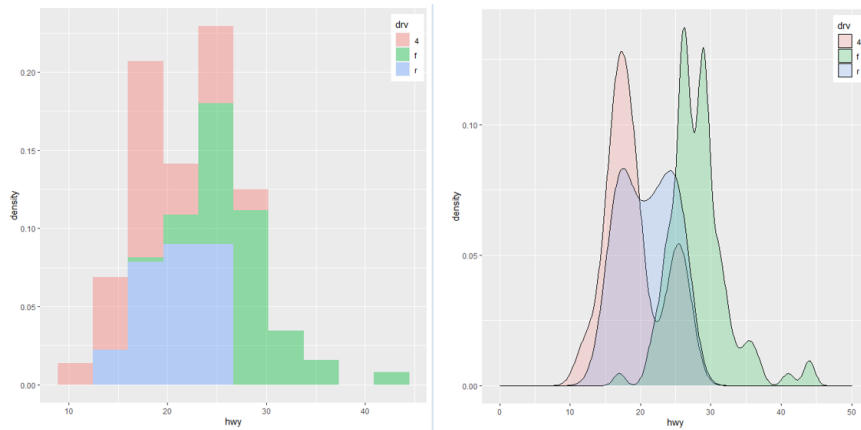
Output 5-22: The overlay of histogram and polygon for count.

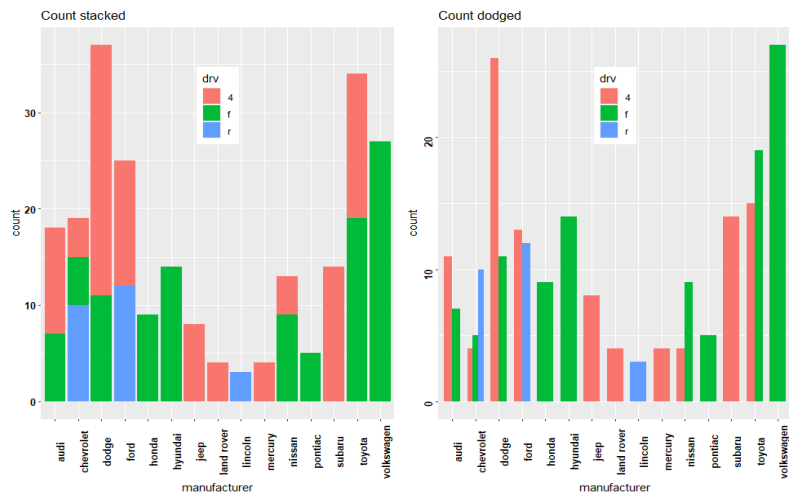**Output 5-23: Histogram subset on fill = drive and by facet wrap on drive.**
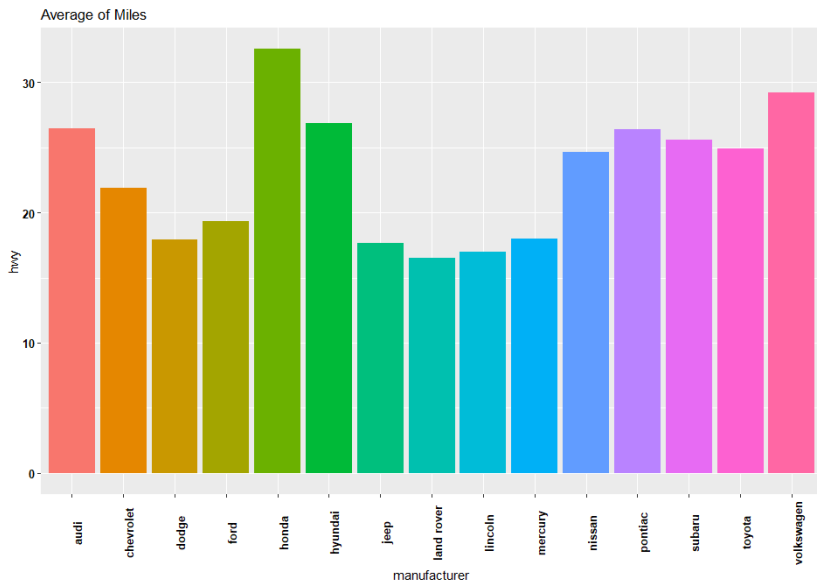
Output 5-24: Polygon alternative of a histogram to compare subsets.

Output 5-25: Density perspective to the histogram and polygon charts of the hwy variable.

Output 5-26: Bar charts of counts and a functions (average).

Output 5-27: Bar chart of mean mpg for manufacturer.

Output 5-28: Line charts of unemployment rate and median length of time unemployed.
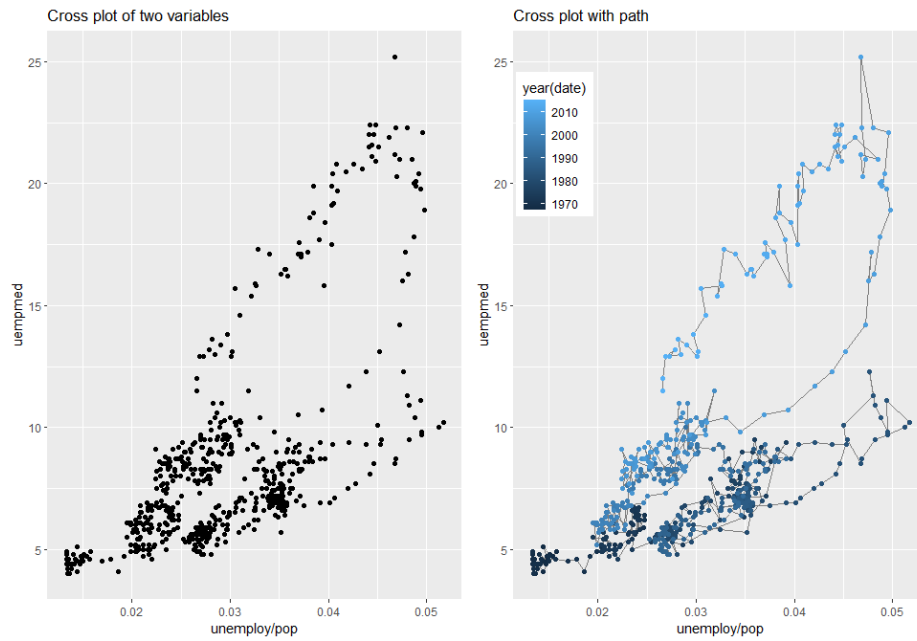
Output 5-29: Scatter plot graph of unemployment and length of unemployment with path line added to show time sequence.
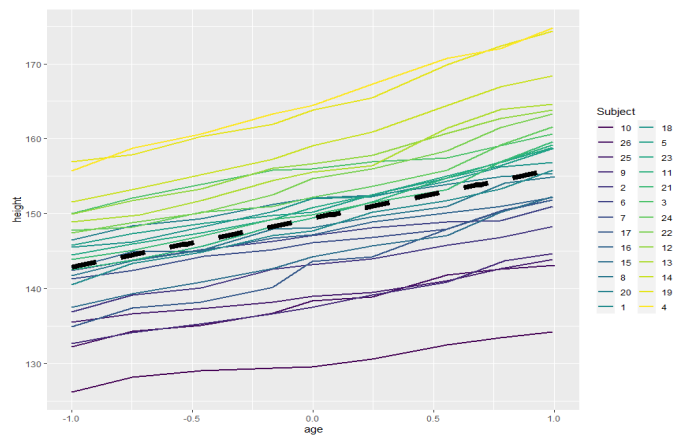
Output 5-30: Lines for groups and linear fit line of the overall data.

Let's look for new lessons in the code to return the output. The code is as follows:
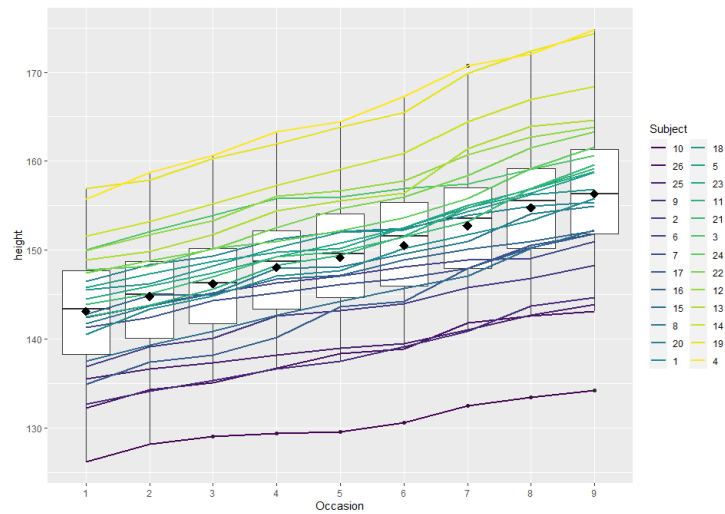
Output 5-31: Centrality and spread of height for each subject at each occasion and individual trends.
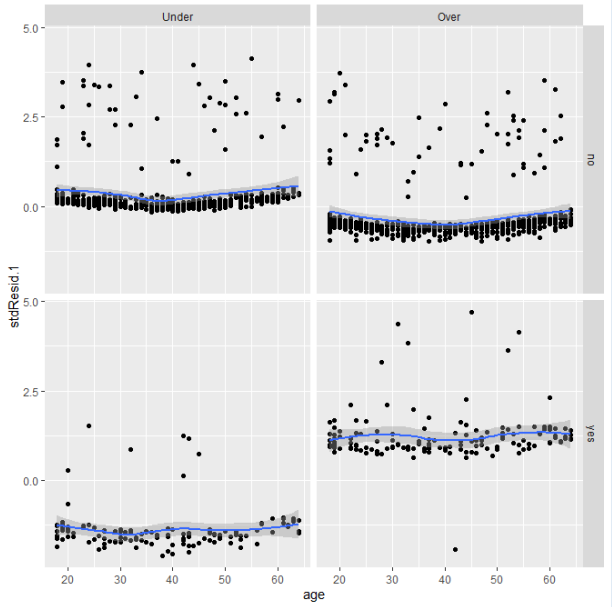
Output 7-10: Check for linearity of the age variable subset on smoker and bmiObese.
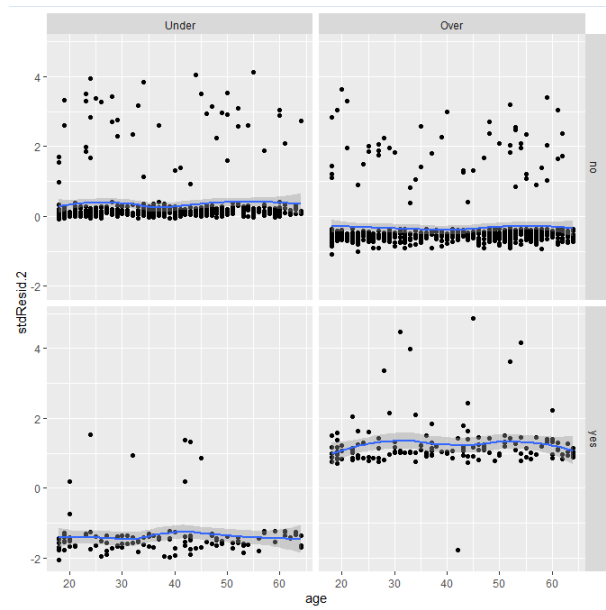
Output 7-11: Linearity is achieved by adding the age squared variable to the model.
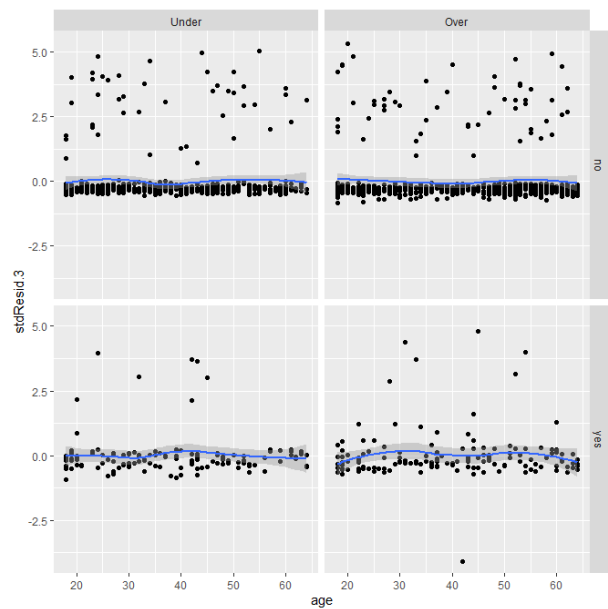
**Output 7-13: Residual plot for age now centered on
residuals with mean equal zero.**
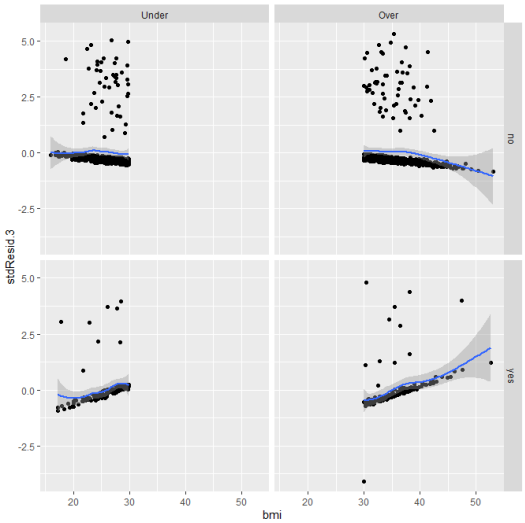
Output 7-14: Residuals to bmi show linearity, but outliers pull the plot northward.
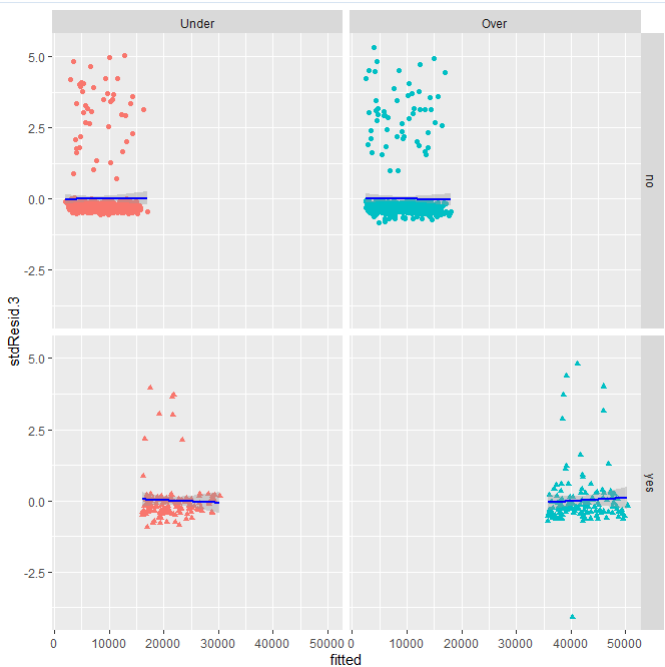
Output 7-17: Expenses, as fitted variable, against standard residuals to model.6.
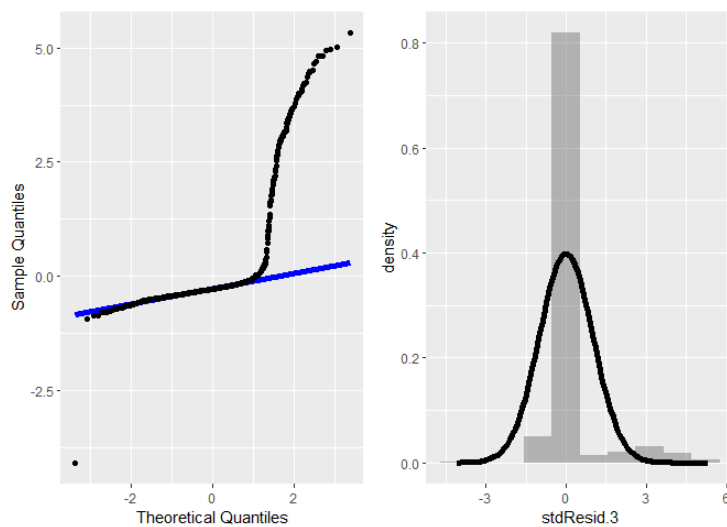
Output 7-19: Q-Q and density plots to inspect for normal distribution.
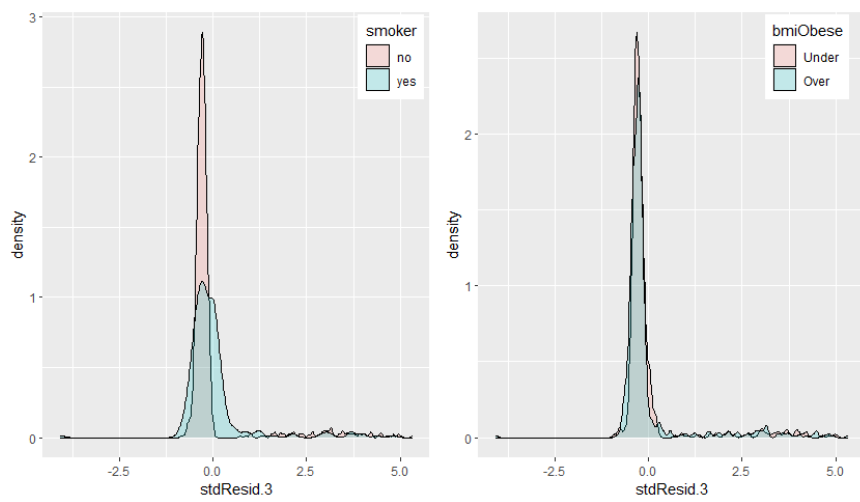
Output 7-20: The distributions are distinctive for smokers but not for bmiObese.
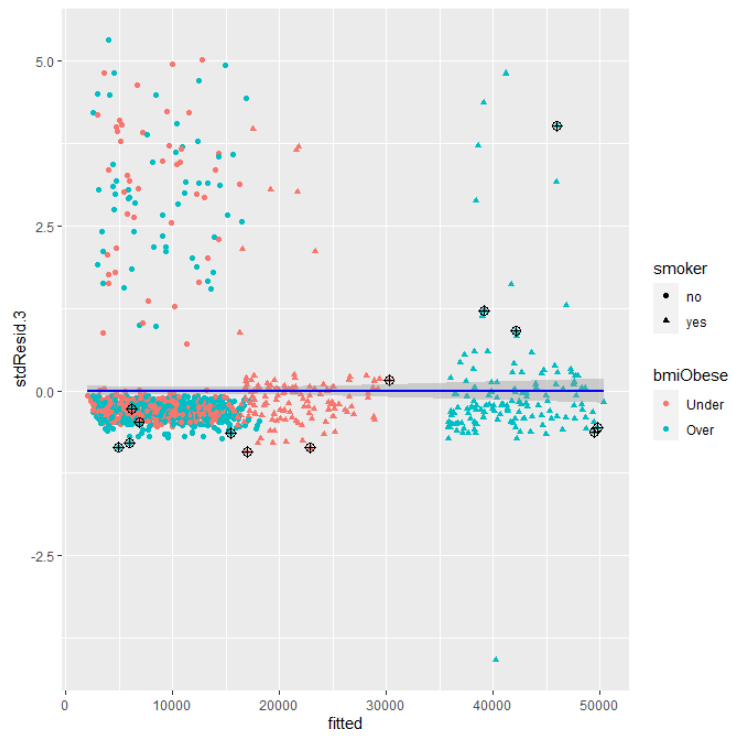
Output 7-21: Leveraging and influencing cases shown among all points.
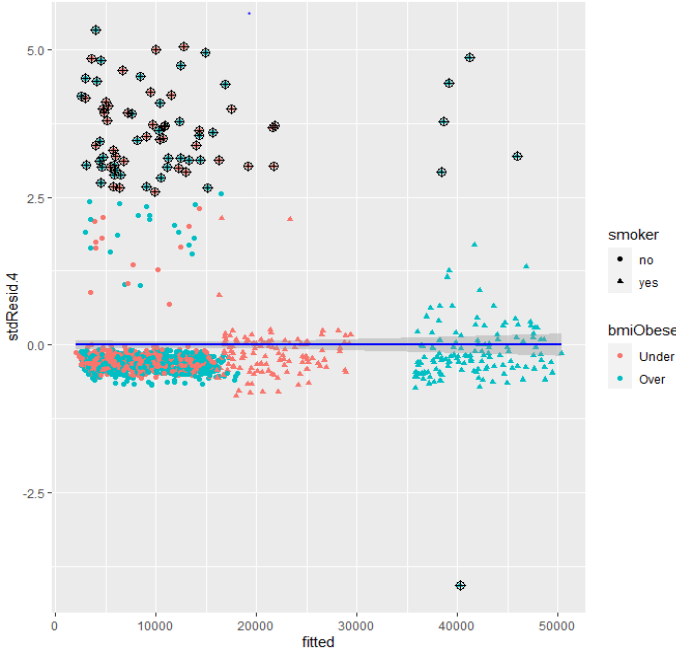
Output 7-22: Outliers cases to the mainstream with 99 percent confidence.
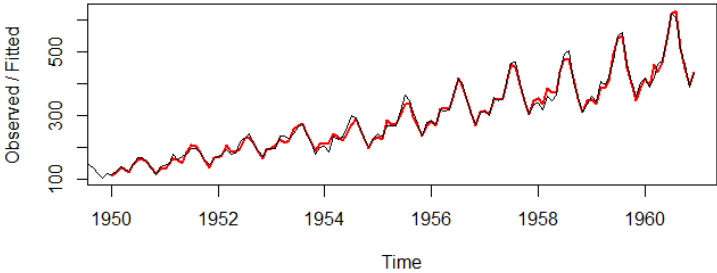
Output 10-20. xhat (red) overlaid on the observations of AirPassenger data.
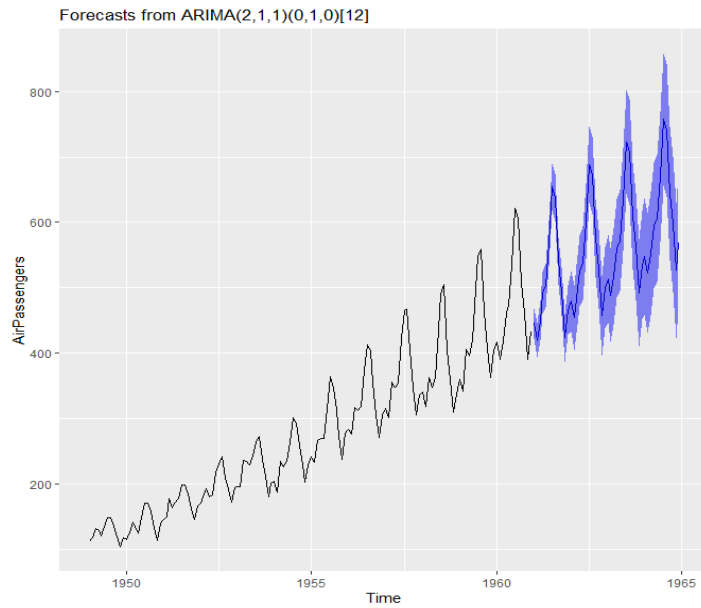
Output 10-23. Past forecast into the future with the ARIMA model.
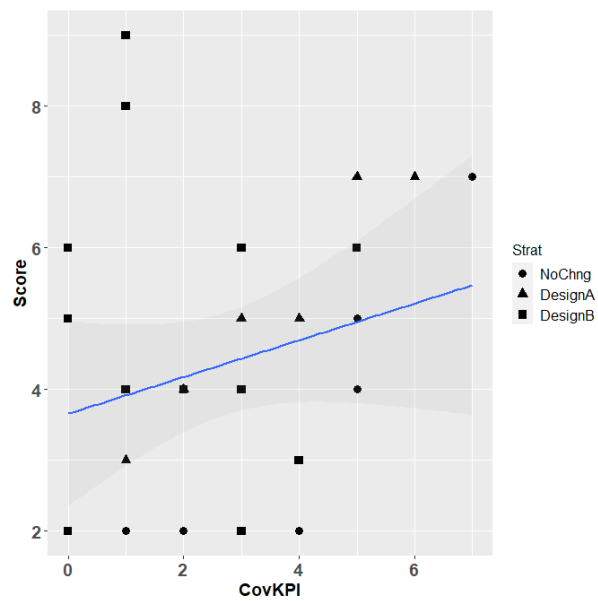
Output 13-22: Scatter plot between Score and CovKPI.
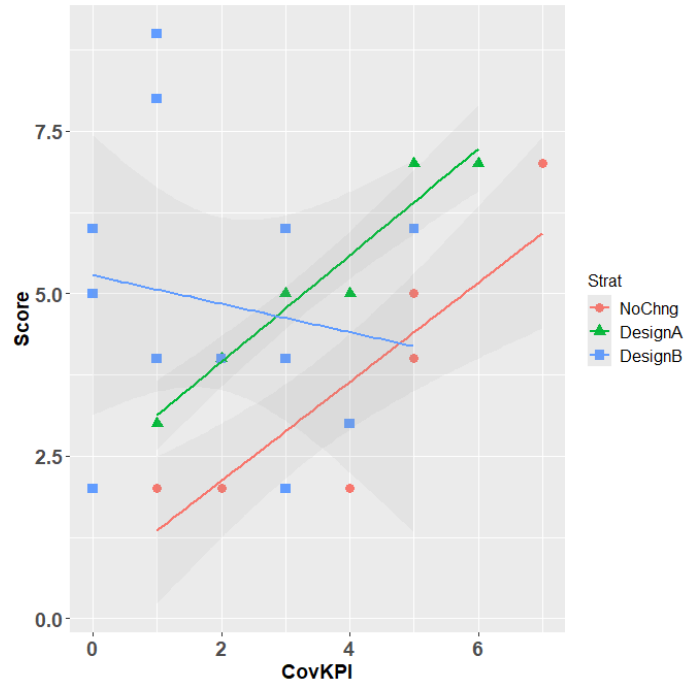
Output 13-25: Slope of regression for the groups.
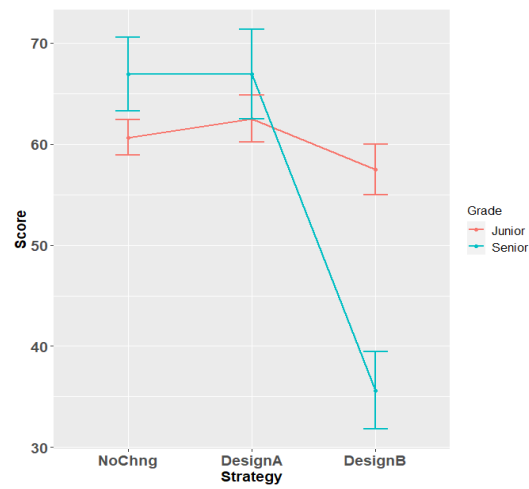
Output 13-30: Plot showing interactions in the ANOVA.

Data and Analytics Skills for Your Career Security
Keeping it simple, only the skills you're likely to need
Richard G. Lamb

For those of us who are role holders in enterprise functioning, the personal purpose of acquiring practical working skills in data and analytics is to be able to better do what we already do and find new ways to do better yet. It follows that if you are a role holder who brings and incorporates data and analytics methods in your thoughts and tasks, your career outlook will be more secure and exciting. The book is written to be your gateway to the skills and to be the templates with which you will install the methods in your operational roles.

We all know that the field of data and analytics is huge and intimidating. It is a long slog to becoming comfortable. During the author's own long slog until arriving at the book, something exciting bubbled to the surface. There is a big difference between what we need to know and everything there is to know. We need to know what is possible as insight for decisions and functioning, we need to know how to get to the insight and, finally, we need to be able to interpret the insight. Just as the book does, we can leave the rest to the data scientists.

About the Author: In 2003, Richard Lamb, while struggling to get at the history captured in the databases of operational systems, found the skills to extract datasets of related history and join them in a super table of variables to make possible what was being envisioned for operational effectiveness. In 2014, Richard realized that, with statistical analytics and free enabling powerful pc-level software, an enterprise could ask and answer questions of operational effectiveness that are otherwise not possible. His activism to bring the epiphanies into the careers of role holders in the mainstream of operations has arrived at this book to explain data and analytics through the demonstration of methods.

Richard is a Registered Professional Engineer and Certified Public Accountant. He has previously authored two books: Availability Engineering and Management for Manufacturing Plant Performance, and Maintenance Reinvented for Business Performance. He has a BSCE, BBA and MBA from the University of Houston and a graduate-level Applied Statistics Certificate from the Texas A&M University.

https://analytics4strategy.com/data-and-code

Analytics4Strategy