

STK 353

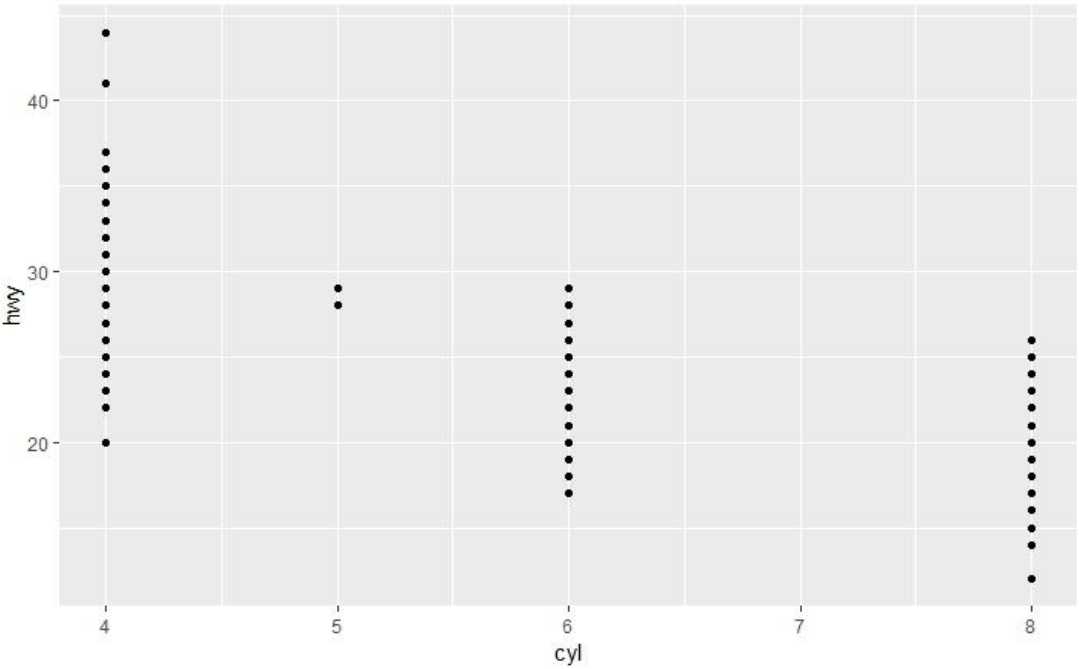
Practical 1: Basics

Name: SJP Eloff

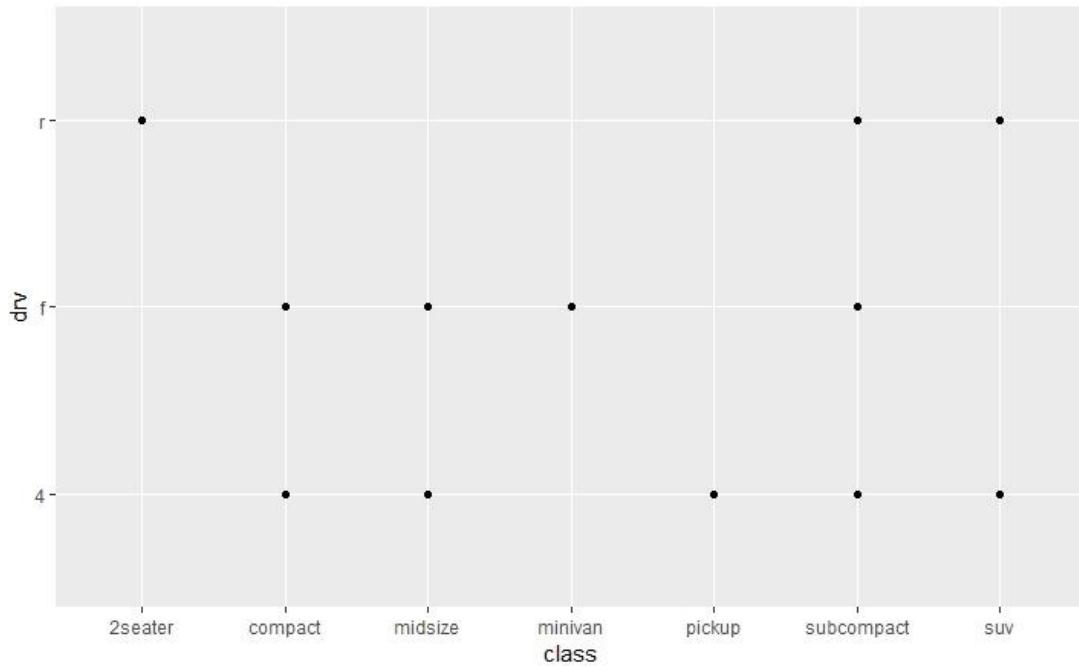
Student number: 10237161

Submission: 23 July 2018, 17:30

Answer Sheet

1	A grey box shows up under the 'Plot' tab, where I assumed the plotted graph was supposed to appear.												
2	The dataset <i>mpg</i> consists of 234 rows and 11 columns.												
3	The variable named <i>drv</i> describes to which wheels the vehicle's power output is sent.												
4	<p>Plot: <i>hwy</i> vs. <i>cyl</i></p>  <table border="1"><caption>Approximate data points from the 'hwy vs. cyl' plot</caption><thead><tr><th>cyl</th><th>hwy</th></tr></thead><tbody><tr><td>4</td><td>15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45</td></tr><tr><td>5</td><td>28, 29, 30</td></tr><tr><td>6</td><td>17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30</td></tr><tr><td>7</td><td></td></tr><tr><td>8</td><td>12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27</td></tr></tbody></table>	cyl	hwy	4	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45	5	28, 29, 30	6	17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	7		8	12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27
cyl	hwy												
4	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45												
5	28, 29, 30												
6	17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30												
7													
8	12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27												

5

Plot: *drv* vs. *class*

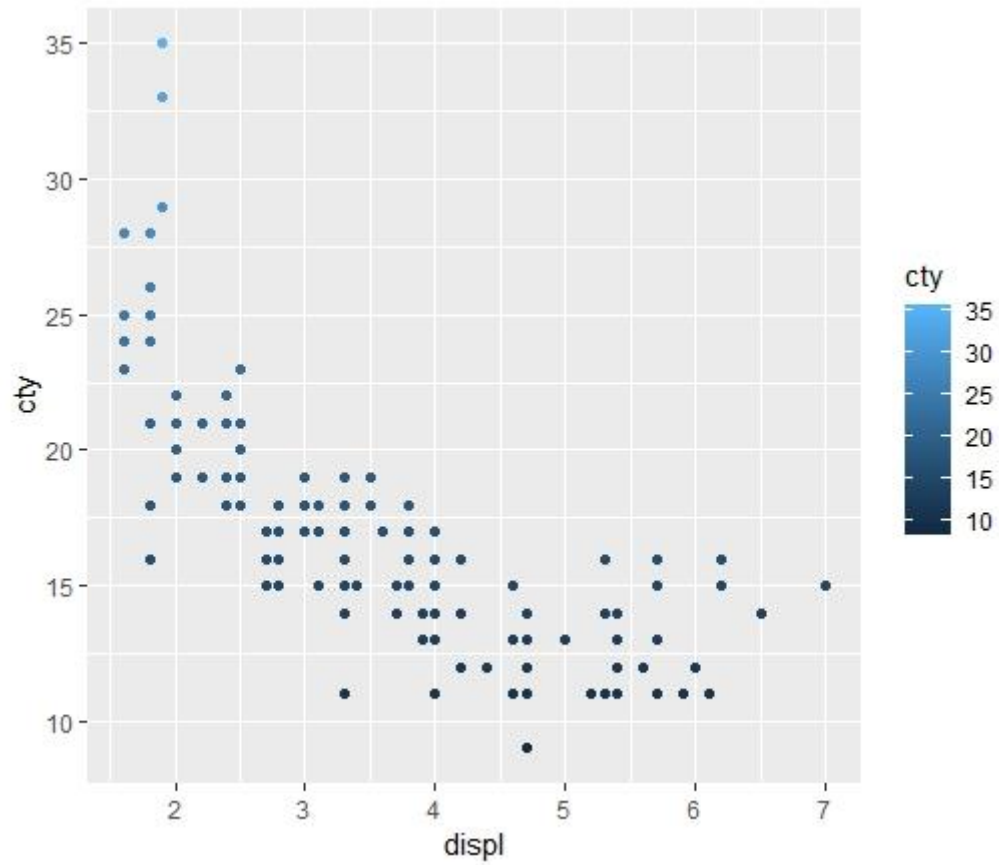
The data doesn't seem to tell us much and is chaotically scattered. This is expected however, since any *drv* can be found in any *class*, i.e. there is no reason to expect that the variables should correlate.

6

Categorical variables: *manufacturer*, *model*, *year*, *trans*, *drv*, *fl*, *class*.Continuous variables: *displ*, *cyl*, *cty*, *hwy*.

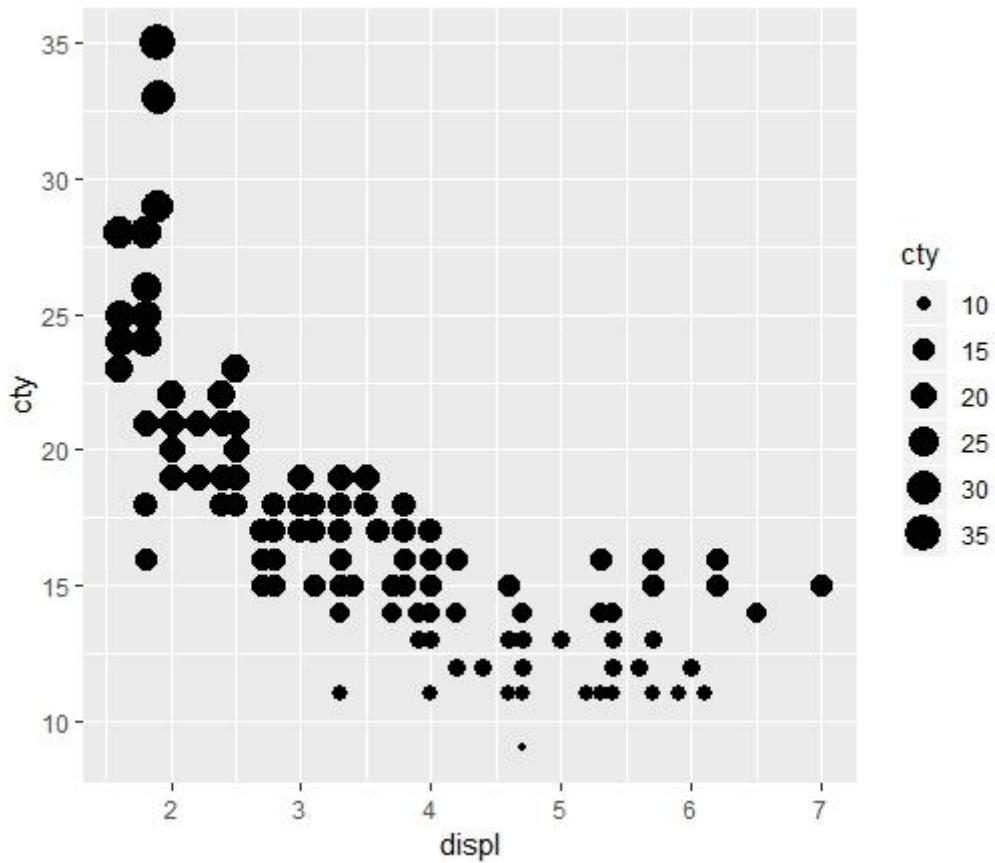
7

Plot: *displ* vs. *cty* (to color)



The color aesthetic maps the magnitude of *cty* to a lighter shade of blue the greater the value.

Plot: *displ* vs. *cty* (to size)



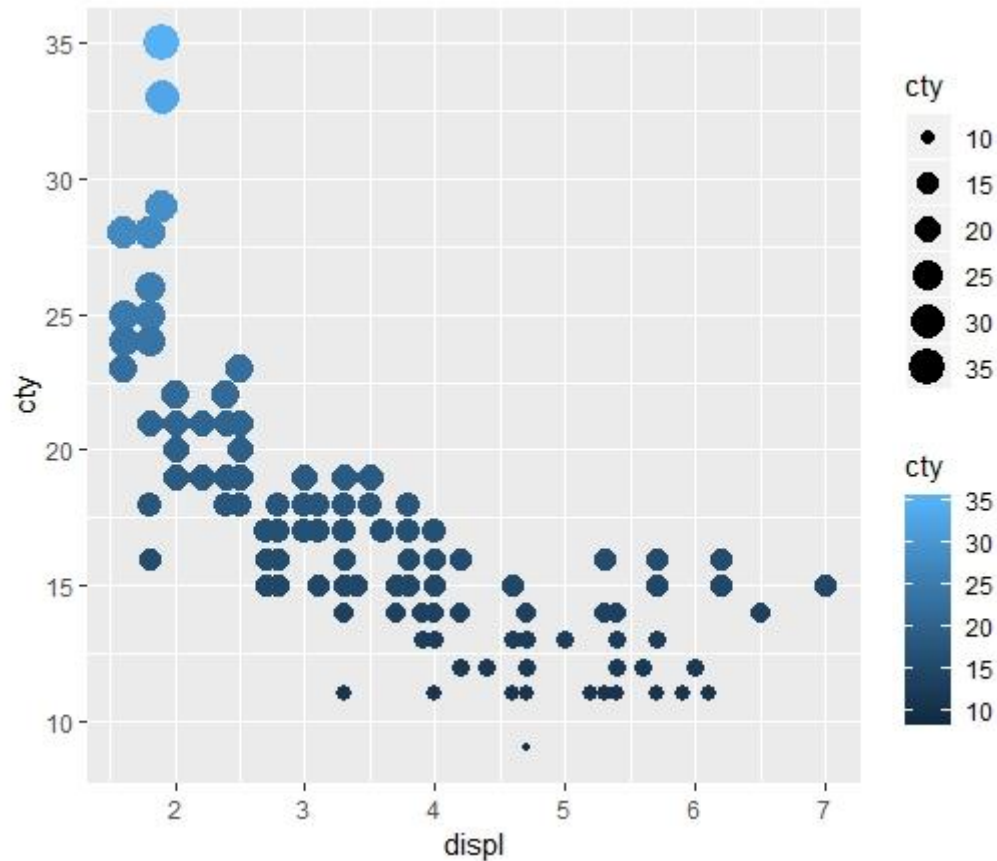
The size aesthetic maps the magnitude of *cty* to a larger black dot the greater the value.

Plot: *displ* vs. *cty* (to shape)

This graph is unattainable since continuous variables cannot be mapped to unordered descriptors e.g. shapes.

8

Plot: *displ* vs. *cty* (to color and shape)



As shown above, multiple aesthetics can be implemented in our graphical representation of the data.

9 The stroke aesthetic magnifies the data points visually. The function `?geom_point` also tells us that the stroke aesthetic defines the width of the border of shapes.

10 To draw a line chart, use `geom_line()`. For a boxplot, use `geom_boxplot()`. For a histogram, use `geom_histogram()`. For an area chart, use `geom_area()`,

Total Marks (out of 10):