**RATHNAVEL SUBRAMANIAM COLLEGE OF ARTS AND SCIENCE (AUTONOMOUS) SULUR, COIMBATORE.**

**SCHOOL OF COMPUTER STUDIES - UG**

**THIRD SEMESTER**

**R FOR DATA SCIENCE**

**ASSIGNMENT QUESTIONS**

**UNIT – I**

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | CO 1 | Remember | What happens if you make a scatterplot of class versus drv? Why is the plot not useful? |
| 2 | CO 1 | Apply | What’s gone wrong with this code? Why are the points not blue?  ggplot(data = mpg) +  geom\_point(mapping = aes(x = displ, y = hwy, color = "blue")) |
| 3 | CO1 | Remember | 1. What happens if you map a aesthetic to something other than a variable name, like aes(colour = displ < 5)? 2. What does the stroke aesthetic do? What shapes does it work with? |
| 4 | CO1 | Apply | Read? Facet\_wrap. What does nrow do? What does ncol do? What other options control the layout of the individual panels? Why doesn’t facet\_grid() have nrow and ncol variables? |
| 5 | CO1 | Understand | 1. What plots does the following code make? What does . do?   ggplot(data = mpg) +  geom\_point(mapping = aes(x = displ, y = hwy)) +  facet\_grid(drv ~ .)  ggplot(data = mpg) +  geom\_point(mapping = aes(x = displ, y = hwy)) +  facet\_grid(. ~ cyl)   1. When using facet\_grid() you should usually put the variable with more unique levels in the columns. Why? |
| 6 | CO1 | Analyze | What geom would you use to draw a line chart? A boxplot? A histogram? An area chart? |
| 7 | CO1 | Analyze | Will these two graphs look different? Why/why not?  ggplot(data = mpg, mapping = aes(x = displ, y = hwy)) +  geom\_point() +  geom\_smooth()  ggplot() +  geom\_point(data = mpg, mapping = aes(x = displ, y = hwy)) +  geom\_smooth(data = mpg, mapping = aes(x = displ, y = hwy)) |
| 8 | CO1 | Remember | What does geom\_col() do? How is different to geom\_bar()? |
| 9 | CO1 | Remember | What is the default geom associated with stat\_summary()? How could you rewrite the previous plot to use that geom function instead of the stat function? |
| 10 | CO1 | Analyze | What does the plot below tell you about the relationship between city and highway mpg? Why is coord\_fixed() important? What does geom\_abline() do?  ggplot(data = mpg, mapping =aes(x =cty, y =hwy)) +  geom\_point() +  geom\_abline() +  coord\_fixed() |