Problem Solving Tasks 2

Your name

3/22/2019

Problem solving task 2 very brief solutions and marking scheme (not for release to students)

Below are just suggestions for plots that answer the questions.

The total workshop is worth 6%, and there are 7 questions. Please allocate marks as follows:

0.5 marks per question for choosing an appropriate plot – i.e. one that answers the question posed, uses the correct parameters in the dataset, and matches the type of data (MOST CRITICALLY that they haven't used bar plots for anything other than categorical data!). I have given suggestions below, but these are not the only options. – total 3.5%

2 marks overall for making effective choices when revising the aesthetic elements of the plots. Please consider as examples, but not an exhaustive list:

- Effective use of colour
- Readable and informative labels
- Effective reorganisation (e.g. rotating and sorting bar plots)
- No unnecessary visual clutter, unless clearly justified

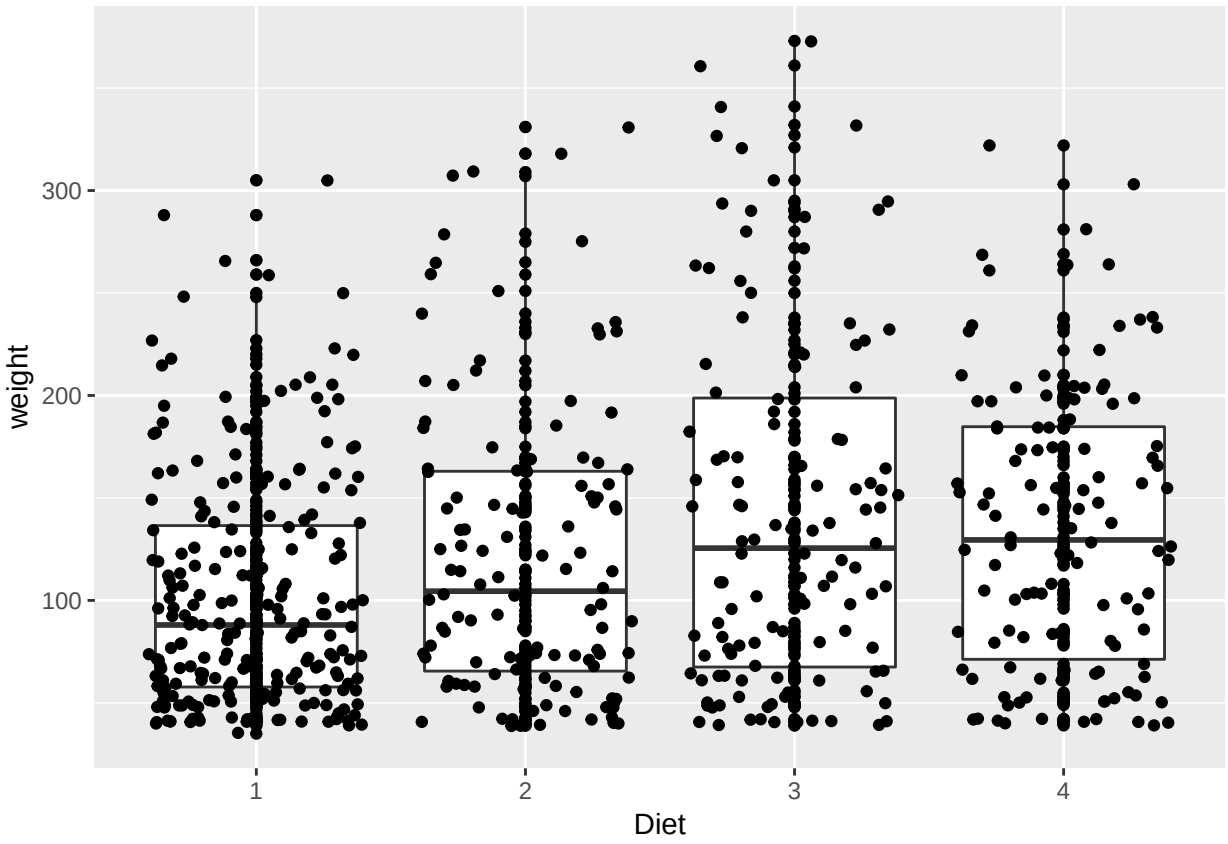
0.5 marks overall for explicitly answering the questions posed – these answers don't have to be correct, but it's not enough just to print the plots with no tie to the question

1 mark overall for well-reasoned justifications and a successfully knitted file.

Question 1

Using the dataset `ChickWeight`, does diet impact weight?

```
ggplot(ChickWeight, aes(x = Diet, y = weight)) + geom_boxplot() + geom_point()  + geom_jitter()
```

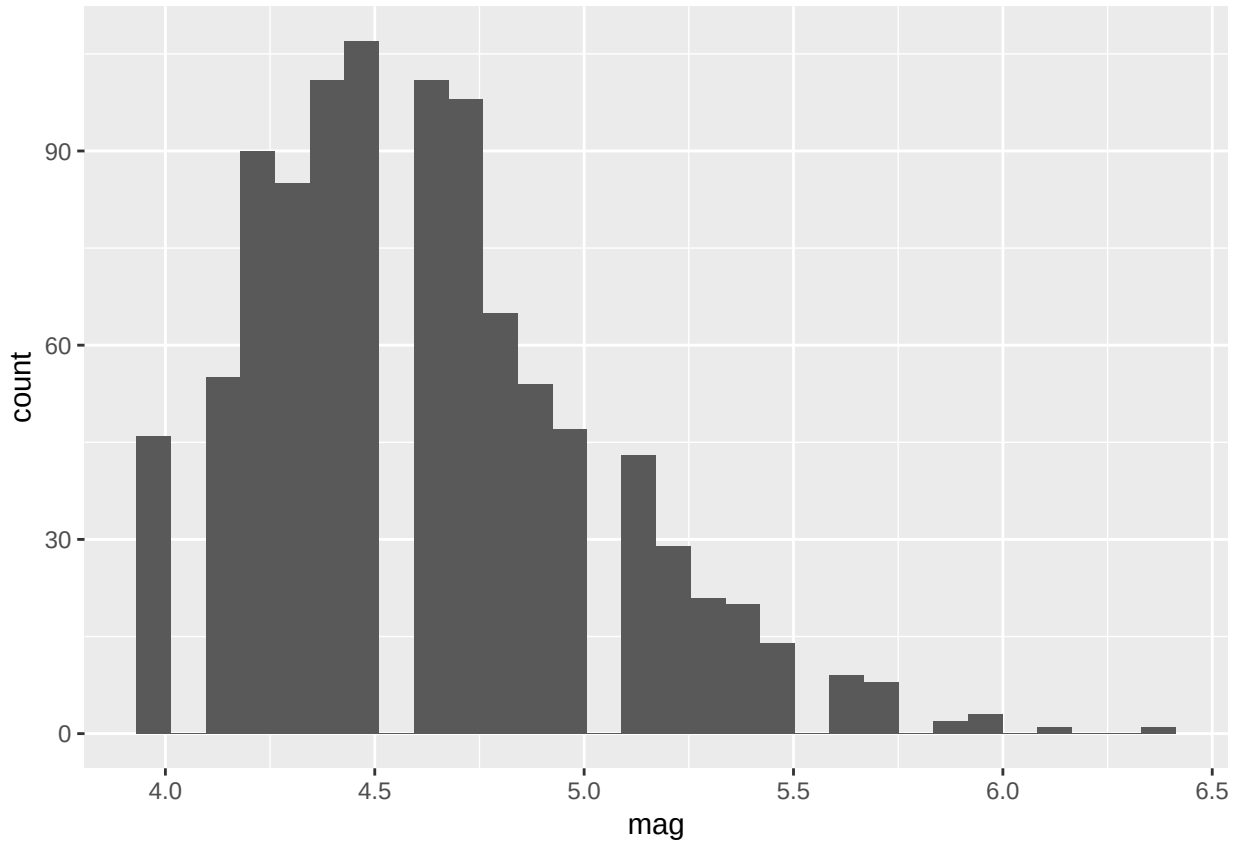


Question 2

Using the dataset `quakes`, design a visualisation that shows the variation of earthquake magnitudes

```
ggplot(quakes, aes(x = mag)) + geom_histogram()
```

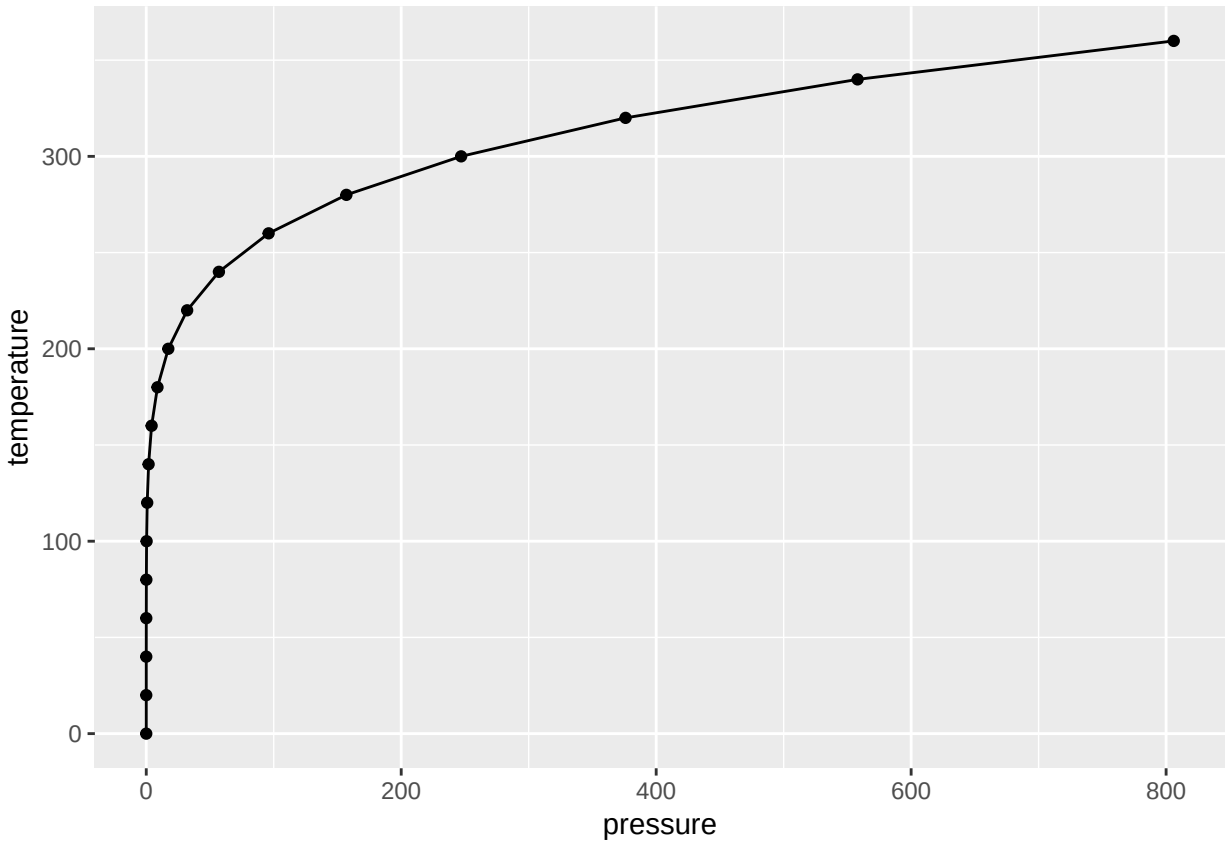
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Question 3

Using the dataset `pressure`, are temperature and pressure related?

```
ggplot(pressure, aes(pressure, temperature)) + geom_point() + geom_line()
```



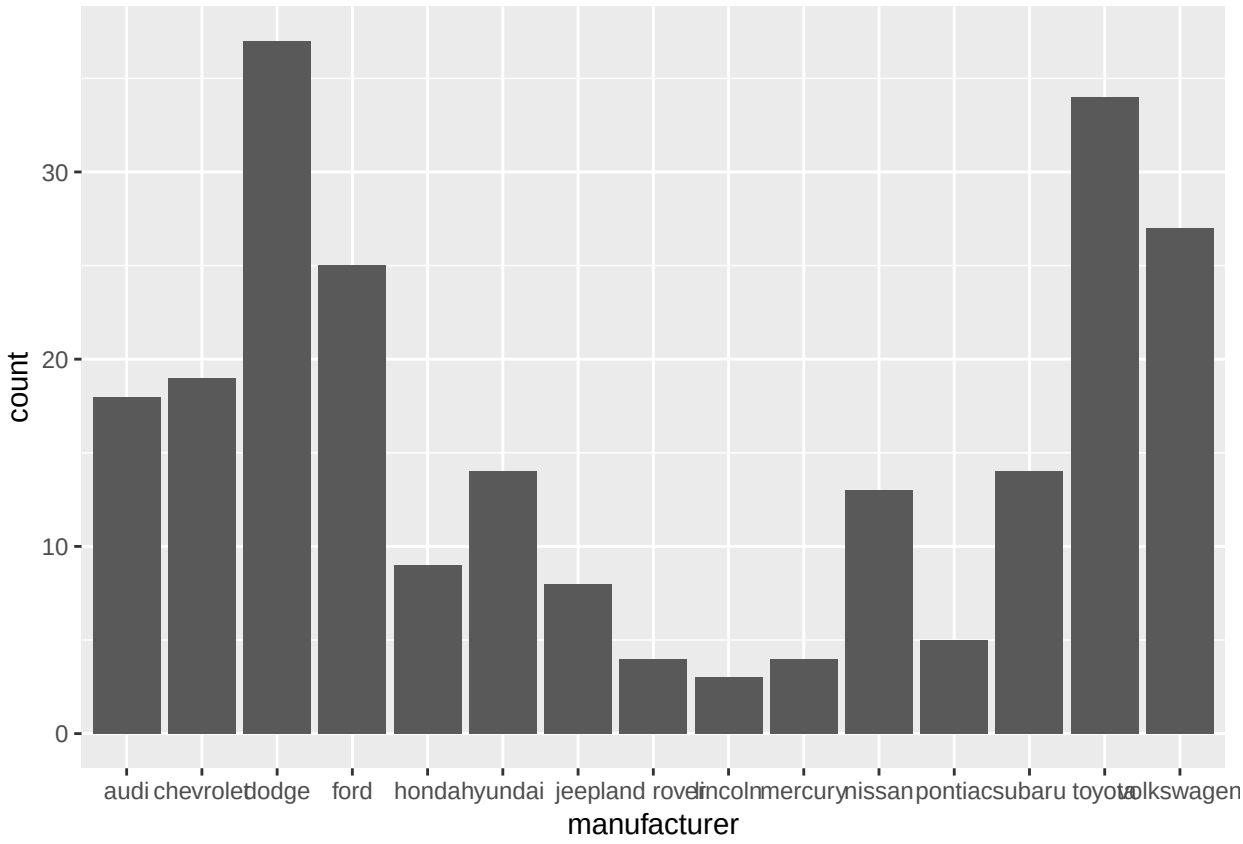
Question 4

Using the dataset `mpg`

4a)

What are the most and least popular car manufacturers?

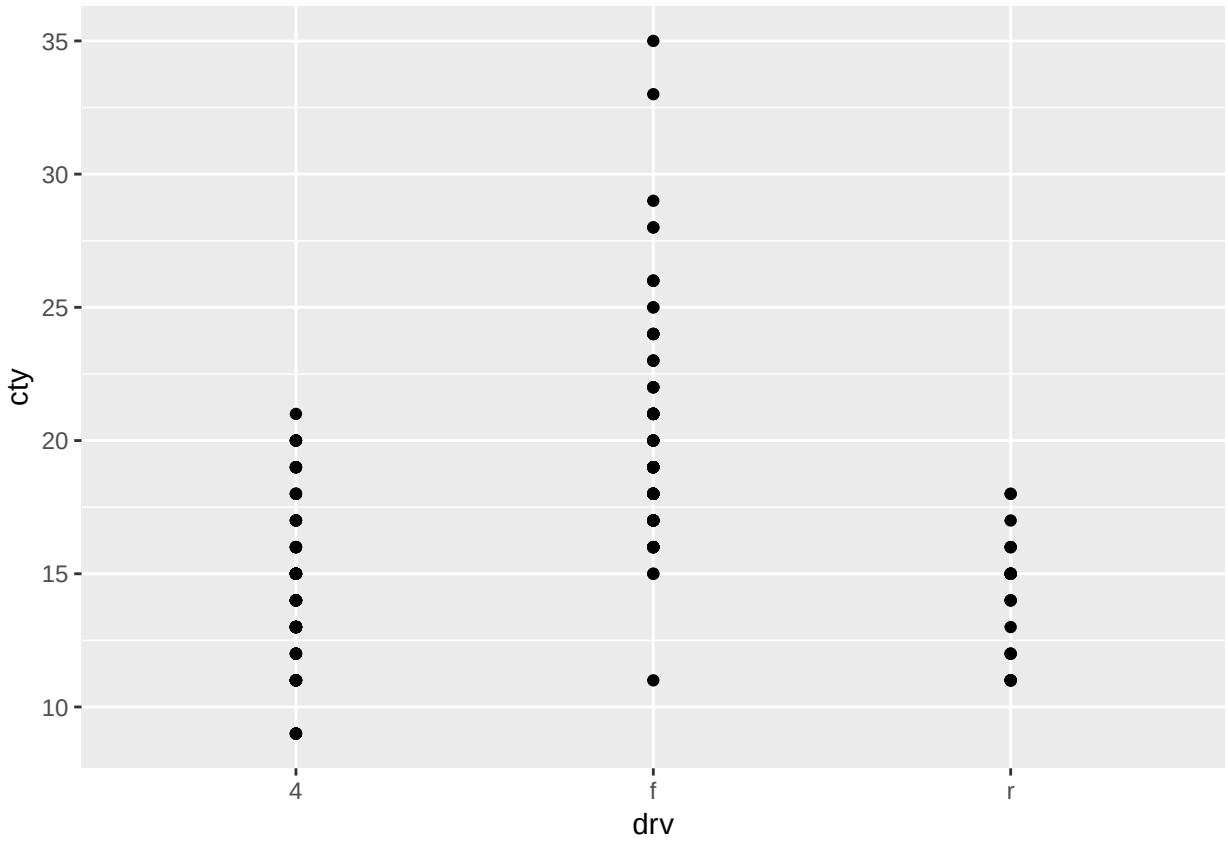
```
ggplot(mpg, aes(manufacturer)) + geom_bar()
```



4b)

Do city miles per gallon (`cty`) correlate or interact with any other variables? Explore 2.

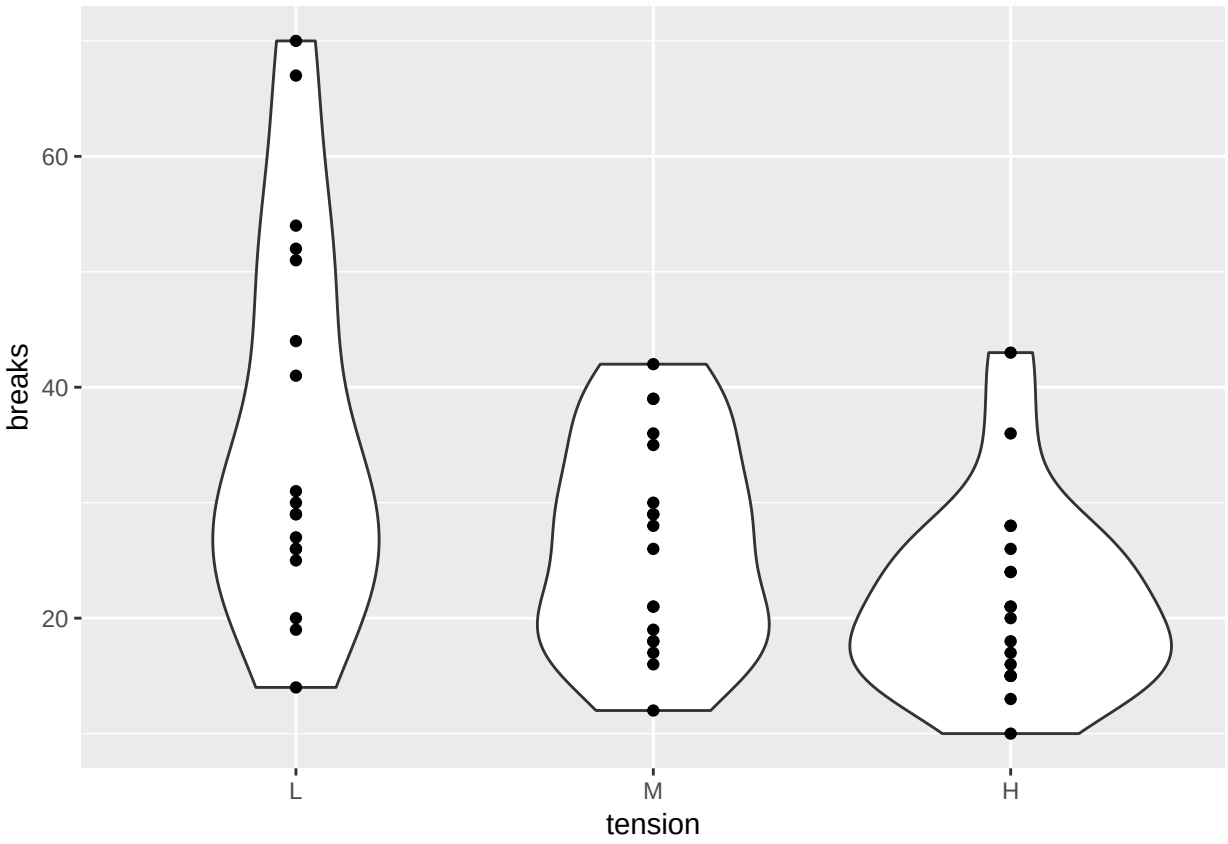
```
ggplot(mpg, aes(drv, cty)) + geom_point()
```



Question 5

Using the dataset `warpbreaks`, what should I do to minimise my yarn breakage?

```
ggplot(warpbreaks, aes(x = tension, y = breaks)) + geom_violin() + geom_point()
```



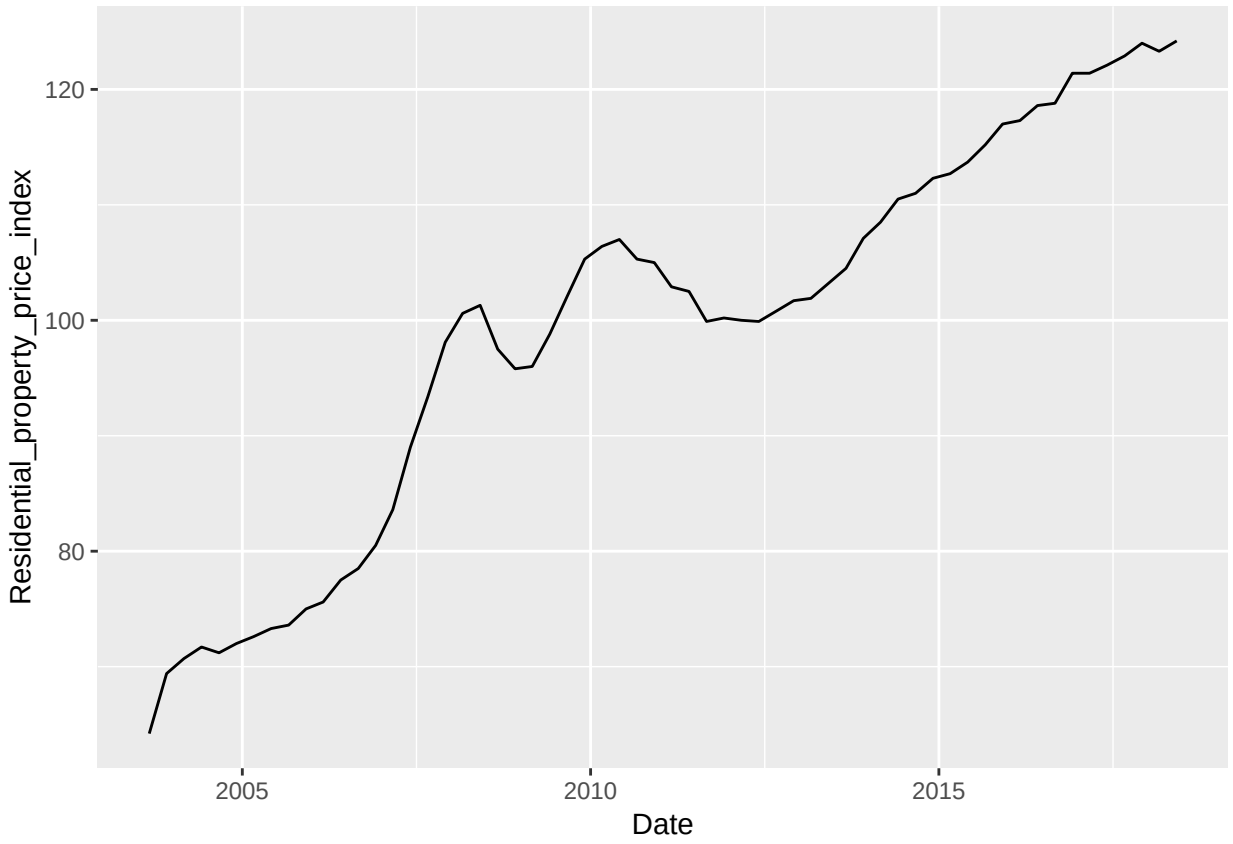
Question 6

6a)

Using the dataset `residential_property_prices` (within the `grattan` package), when should I have bought a house in Brisbane? Give me a few recommendations.

```
x <- residential_property_prices %>%
  filter(City == 'Brisbane')

ggplot(x, aes(Date, Residential_property_price_index)) + geom_line()
```



6b)

If I want to be as cheap as possible, has my best choice changed since 2005?

Question 7

Type `data()` into the console to see a list of all data available in R base or other packages. Explore any dataset you like, produce three plots, and write 1 sentence about what I might discover from that plot