Assignment 3. ST661 2018
Catherine Hurley
Due on Wednesday November 21 6pm.

You will work in pairs (see attached lists).
You should complete this assignment using Rmarkdown. Place the printed html file (with both names) in the box labelled ST661/ST663 in the ground floor of Logic house (under stairs).
Also one person in each pair should upload the .Rmd file to Moodle.

1. Type in

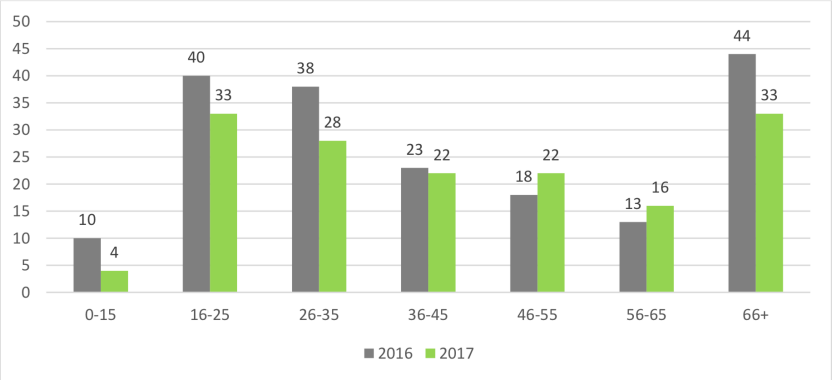
```
library(ggplot2)
library(MASS)
head(Pima.tr)
```

This code loads the data set `Pima.tr`, and its help file. Use ggplot2 for all plots. Using this dataset:

   (a) Make a scatterplot plot of `bp` versus `npreg`.
   (b) Using the function `cut_interval`, construct a factor version of `npreg` with n=4 levels. Call this new variable `npregf`. Add this variable to dataset `Pima.tr`.
   (c) Plot boxplots of `bp` for each `npregf` level.
   (d) Make a scatterplot of `glu` versus `age`. Use colour to show variable `type` and add smooths for the two groups.
   (e) Redo the previous plot, separating out the two types. (Colour is not now needed)

2. The Irish road safety authority recently produced a report with the graph below. Reconstruct it using ggplot. (Use google for help!)

Figure 8. Deaths by age group, January to December 31st 2017 vs 2016



3. The data Hep2012.csv contains data with the points athletes received from each event in the Heptathlon in the 2012 Olympics. Access the data with

```
hep <- read.csv("Hep2012.csv")
```

You might also find this 'long' version of the data useful:

```
library(tidyverse)
heplong <- gather(hep, key=Event, value=Points,-Athlete)
```

You should

   (a) Pose a question relating to the data.
   (b) Using ggplot, construct a visualisation which helps to answer this question.
   (c) Answer your question, based on the visualisation.
   (d) Explain why you selected this particular visualisation.

Your answer should fit on one page.

Pairs for assignment 3.

First name	Surname	Email address	
SHUBHAM	AGGARWAL	SHUBHAM.AGGARWAL.2019@MUMAIL.IE	1
IAN	FINNEGAN	IAN.FINNEGAN.2019@MUMAIL.IE	1
ROSS	AIROLDI	ROSS.AIROLDI.2019@MUMAIL.IE	2
SEAN	O'RIOGAIN	SEAN.ORIOGAIN.2019@MUMAIL.IE	2
VINEET	BHATNAGAR	VINEET.BHATNAGAR.2019@MUMAIL.IE	3
VINCENT	TENNANTY	VINCENT.TENNANTY.2019@MUMAIL.IE	3
PRATIK KUMAR	BUDHDEO	PRATIK.BUDHDEO.2019@MUMAIL.IE	4
CATHERINE ELIZABETH	DENNISON	catherine.dennison.2015@mumail.ie	4
AMIT	CHINWAN	AMIT.CHINWAN.2019@MUMAIL.IE	5
MICHAEL	EVISTON	MICHAEL.EVISTON.2019@MUMAIL.IE	5
DEVAHARSHA	DADI	DEVAHARSHA.DADI.2019@MUMAIL.IE	6
ADAM	DOYLE	ADAM.DOYLE.2019@MUMAIL.IE	6
BHARVI	DHALL	BHARVI.DHALL.2019@MUMAIL.IE	7
MEGAN	WESTON	MEGAN.WESTON.2019@MUMAIL.IE	7
UDO CHURCHILL	OKIKE	OKIKE.UDO.2019@MUMAIL.IE	8
RAMGOKUL	SRINIVASAN	RAMGOKUL.SRINIVASAN.2019@MUMAIL.IE	8
KATHRYN	DILLON	KATHRYN.DUFFY.2019@MUMAIL.IE	9
ADARSHA	HEBBUR GURU CHANNABASAVANNA	ADARSHA.HEBBURGURUCHANNABA.2019@MUMAIL.IE	9
LUKE TREVOR	BROWN	luke.brown.2015@mumail.ie	10
SUCHISMIT	GUPTA	SUCHISMIT.GUPTA.2019@MUMAIL.IE	10
VIVIENNE CATHERINE	DOYLE	vivienne.doyle.2016@mumail.ie	11
SNEHAL UMESH	KADAM	SNEHAL.KADAM.2019@MUMAIL.IE	11
JOHN MARTIN	MC ENEANEY	john.mceneaney.2015@mumail.ie	12
SIMRAN JEET	SINGH	SIMRAN.NA.2019@MUMAIL.IE	12
LAURA FRANCES	MURPHY	LAURA.MURPHY.2019@MUMAIL.IE	13
RISHITA JANARDAN	YADAV	RISHITA.YADAV.2019@MUMAIL.IE	13
RAVINDER PAL	SINGH	RAVINDER.SINGH.2019@MUMAIL.IE	14
PAUL	WILLIAMSON	PAUL.WILLIAMSON.2019@MUMAIL.IE	14
GANESH	SHARMA	GANESH.SHARMA.2019@MUMAIL.IE	15
DAVID	WHITE	DAVID.WHITE.2019@MUMAIL.IE	15
SMEET AJAY	PALKAR	PALKAR.SMEETAJAY.2019@MUMAIL.IE	16
MARK	WILLIAMS	MARK.WILLIAMS.2019@MUMAIL.IE	16
SHANE	HOGAN	SHANE.HOGAN.2019@MUMAIL.IE	17
SAYANTAN	SENGUPTA	SAYANTAN.SENGUPTA.2019@MUMAIL.IE	17

There are many ways to answer this. For example: What distinguishes the top athletes? Do they receive extremely high points for a few events or do they receive good points across the board?