

## Problem 2.5

Stephen Blatti

October 11, 2017

2.5 The data frame vit2005 in the PASWR package contains descriptive info and the appraised total price (in euros) for apartments in Victoria, Spain.

```
library(PASWR2)
```

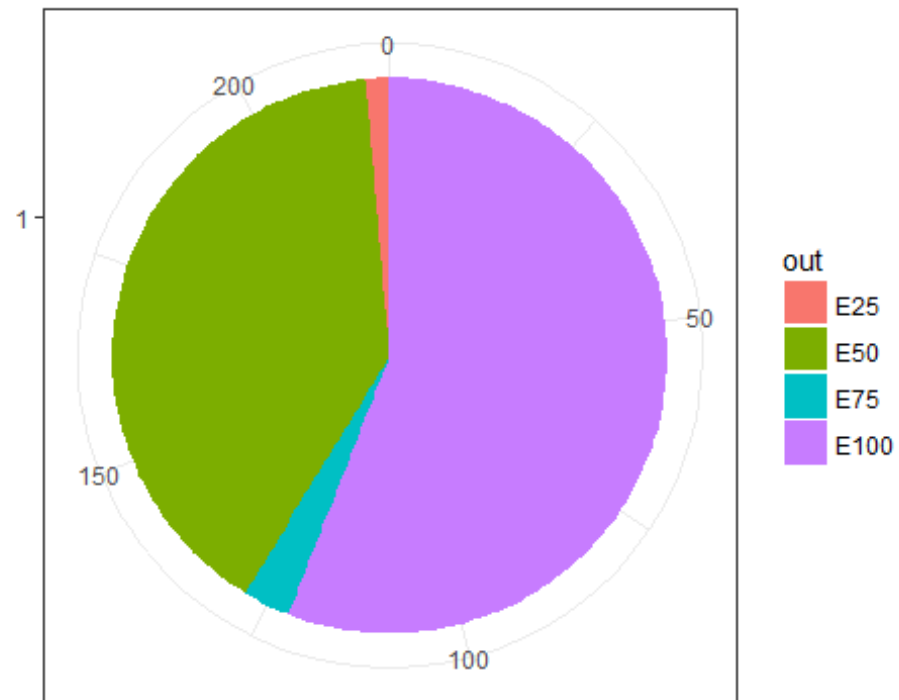
```
## Warning: package 'PASWR2' was built under R version 3.4.2
```

```
## Loading required package: lattice
```

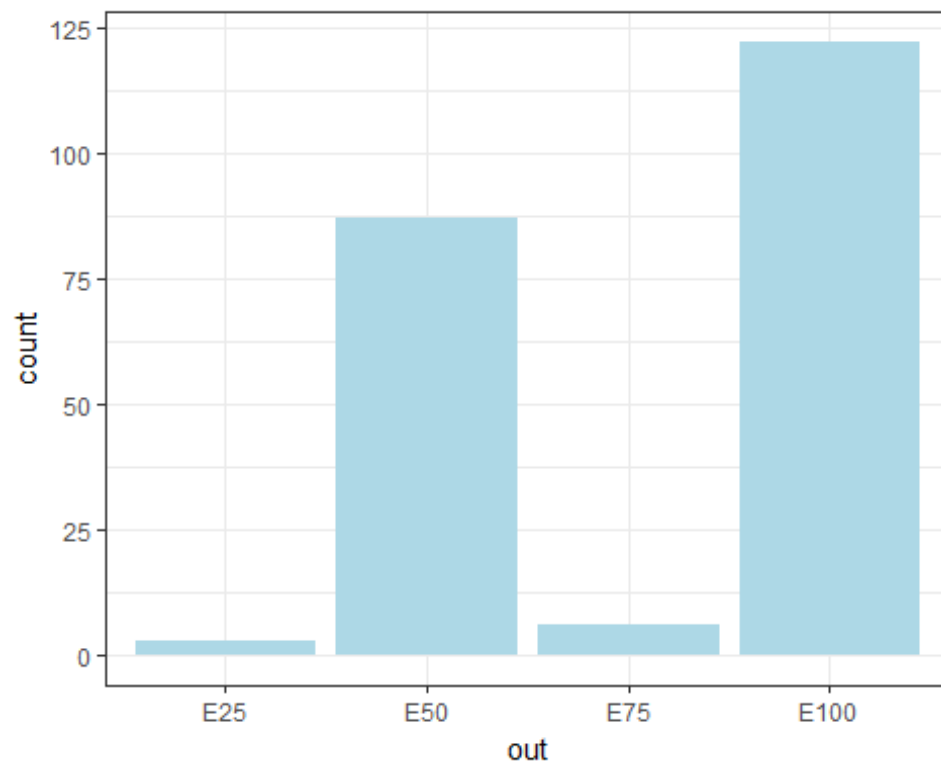
```
## Loading required package: ggplot2
```

a) Create a freq table, piechart, and a barplot showing the num of apartments grouped by the var out. Which method conveys the info best?

```
VIT2005$out <- factor(VIT2005$out, levels = c("E25", "E50", "E75", "E100"))
plot1 <- ggplot(data = VIT2005, aes(x = factor(1), fill = out))+
  geom_bar(width = 1) + coord_polar(theta = "y") + theme_bw() + labs(x = "", y
= "")
plot2 <- ggplot(data = VIT2005, aes(x = out)) + geom_bar(fill = "lightblue")
+ theme_bw()
plot1
```



plot2

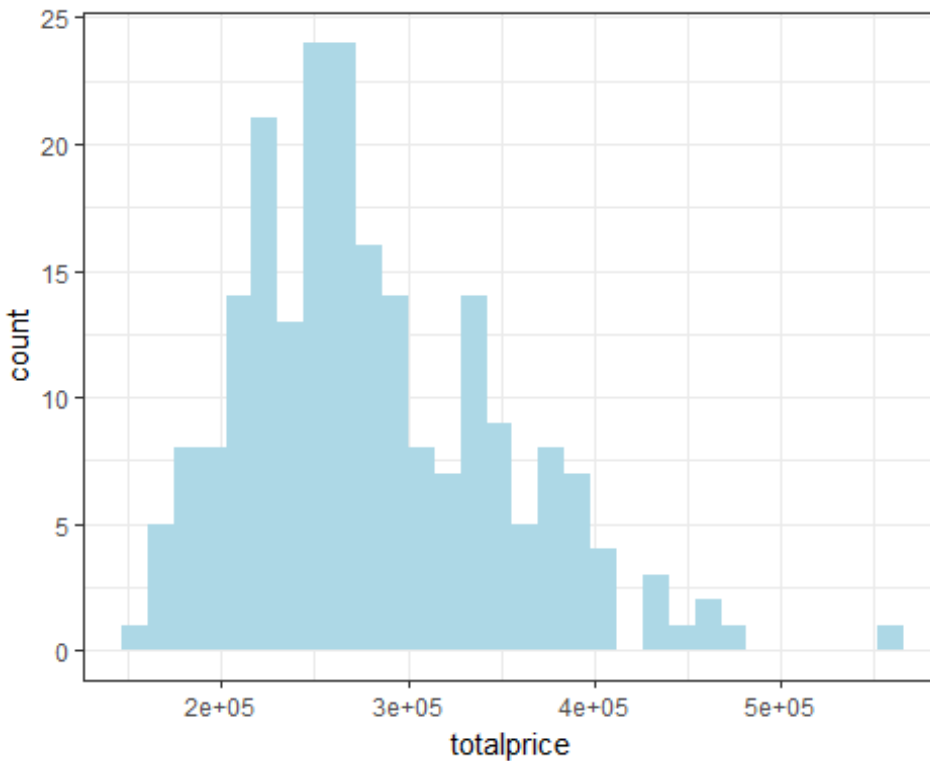


*# The barplot is best*

b) Characterize the distribution of the var totalprice.

```
ggplot(data = VIT2005, aes(x = totalprice)) + geom_histogram(fill = "lightblue") + theme_bw()
```

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



```
max(VIT2005$totalprice) #one outlier at 560000.
```

```
## [1] 560000
```

```
median(VIT2005$totalprice)
```

```
## [1] 269750
```

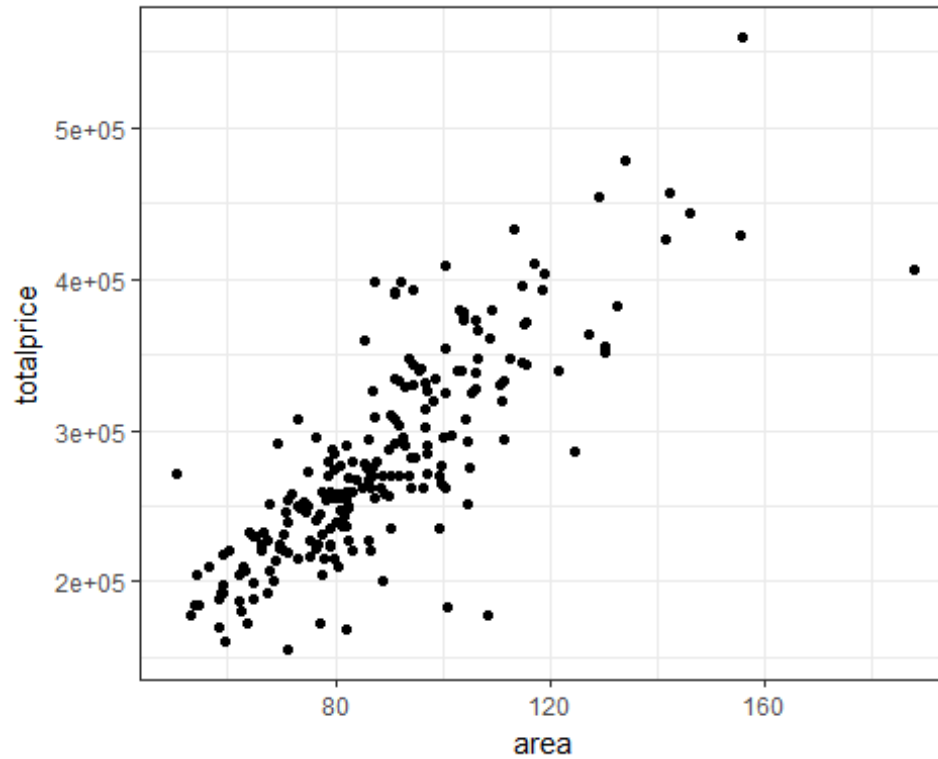
```
IQR(VIT2005$totalprice)
```

```
## [1] 100125
```

*# The distribution of totalprice is skewed to the right. The median totalprice is 269750 and the IQR for totalprice is 100125.*

c) Characterize the relationship between totalprice and area.

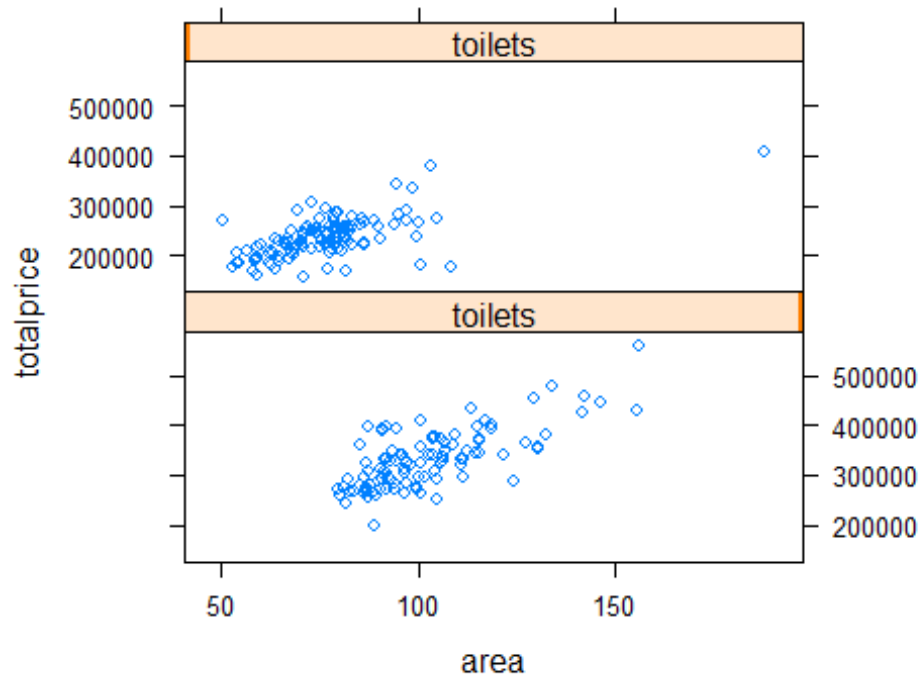
```
ggplot(data = VIT2005, aes(x = area, y = totalprice)) + geom_point() + theme_bw()
```



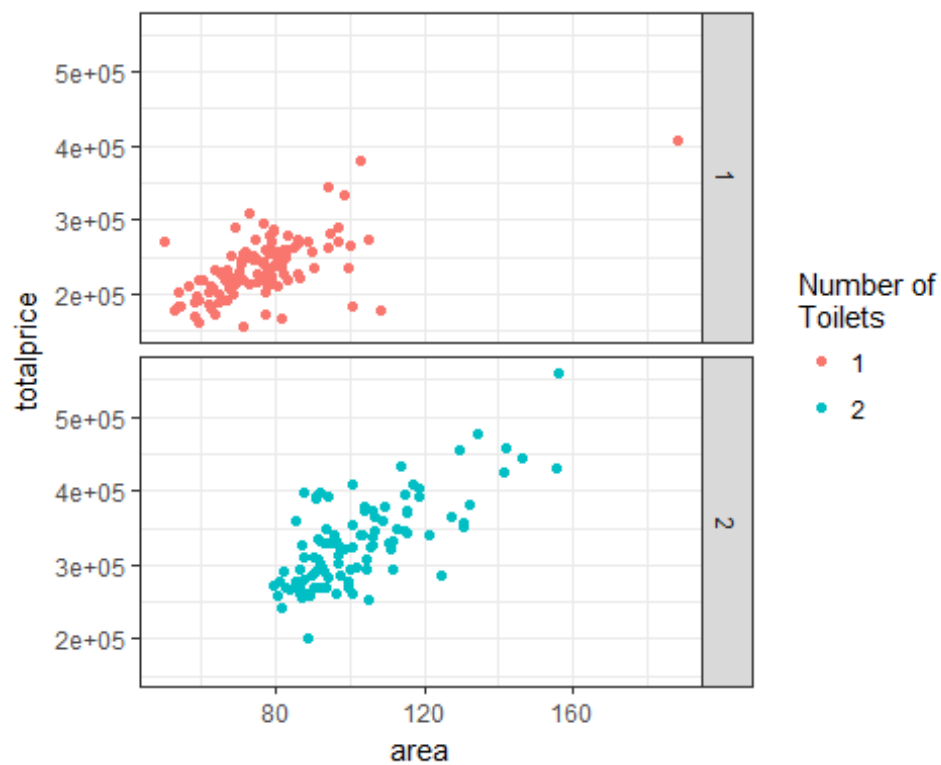
*# The relationship between totalprice and area is of positive linear.*

d) Create a Trellis plot of totalprice vs area conditioning on toilets. Between what 2 values of area do apartments have both one and two bathrooms?

```
xyplot(totalprice ~ area | toilets, data = VIT2005, layout = c(1, 2),  
as.table = TRUE)
```



```
ggplot(data = VIT2005, aes(x = area, y = totalprice, color =
as.factor(toilets))) + geom_point() + facet_grid(toilets ~ .) + theme_bw() +
guides(color = guide_legend("Number of\nToilets"))
```



```
# Apartments with one bathroom: approx between 50 and 100 m^2  
# Apartments with two bathrooms: approx between 80 and 120 m2.  
# The intersection of apartments with one and two bathrooms is approximately  
(80 to 100)m^2.
```

- e) Use the area values from part (d) to create a subset of apartments that have both 1 and 2 bathrooms. By how much does an additional bathroom increase the appraised value of the apartment? Would you be willing to pay for an additional bathroom if you lived in Spain?

```
both_bathrooms <- subset(VIT2005, subset = area >= 80 & area <= 100)  
result <- tapply(both_bathrooms$totalprice, both_bathrooms$toilets, median)  
result
```

```
##      1      2  
## 255000 291000
```

```
diff(result)
```

```
##      2  
## 36000
```

```
# The increase in totalprice for a second bathroom for apartments between  
# 80 and 100 m^2 is 36000 euros - probably not willing to pay the additional  
costs.
```