Wecome to PSYCH 420

Your Name

2025-08-28

This is a heading

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.2     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.1.0
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
```

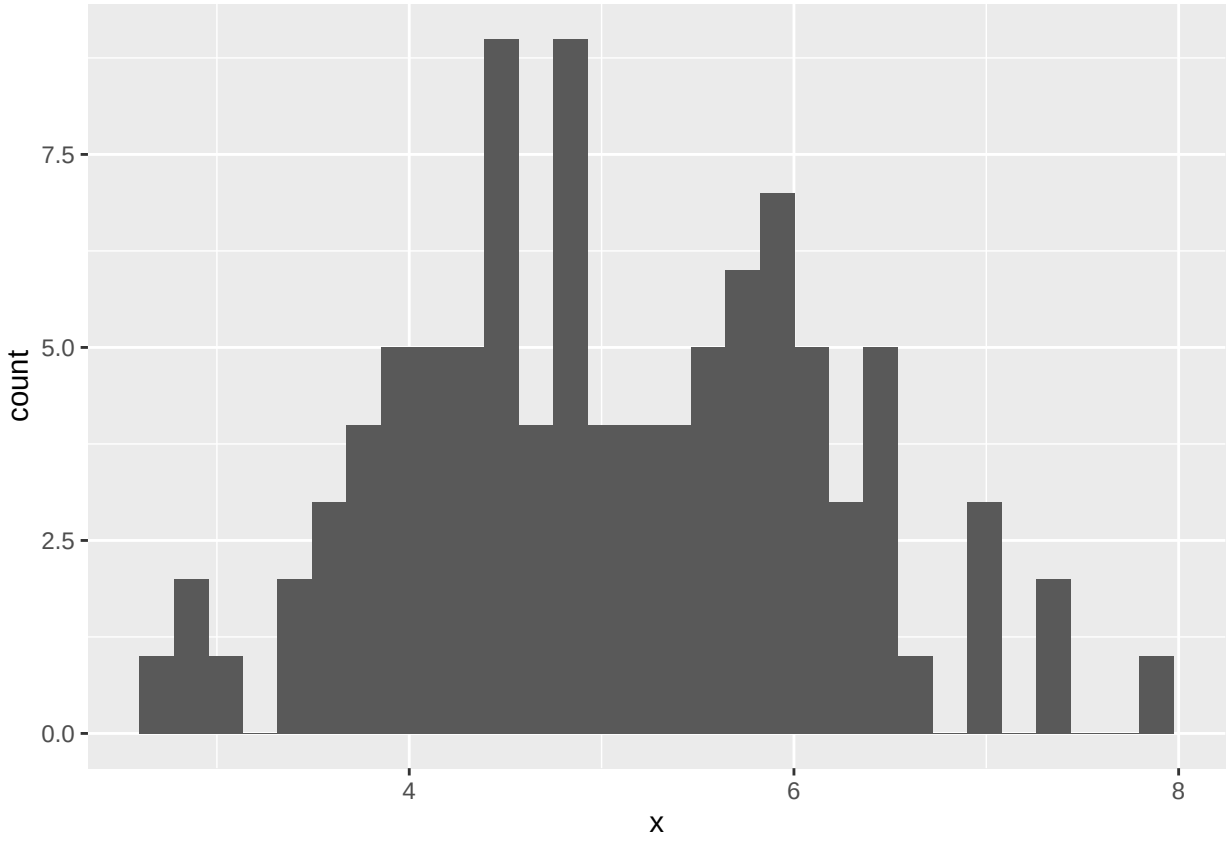
```
a <- 2

b <- 1

x = rnorm(100, mean = 5, sd=1)

errors <- rnorm(100, mean = , sd=1)

data <- tibble(x = x, y = a*x+b + errors)

head(data) # We can take a peak at the first few rows using the head() functions
```

```
## # A tibble: 6 x 2
##       x     y
##   <dbl> <dbl>
## 1  5.84 12.3
## 2  4.23  9.46
## 3  5.72 10.8
## 4  6.33 12.6
## 5  5.33 11.6
## 6  4.56  9.92
```

This is the histogram of x

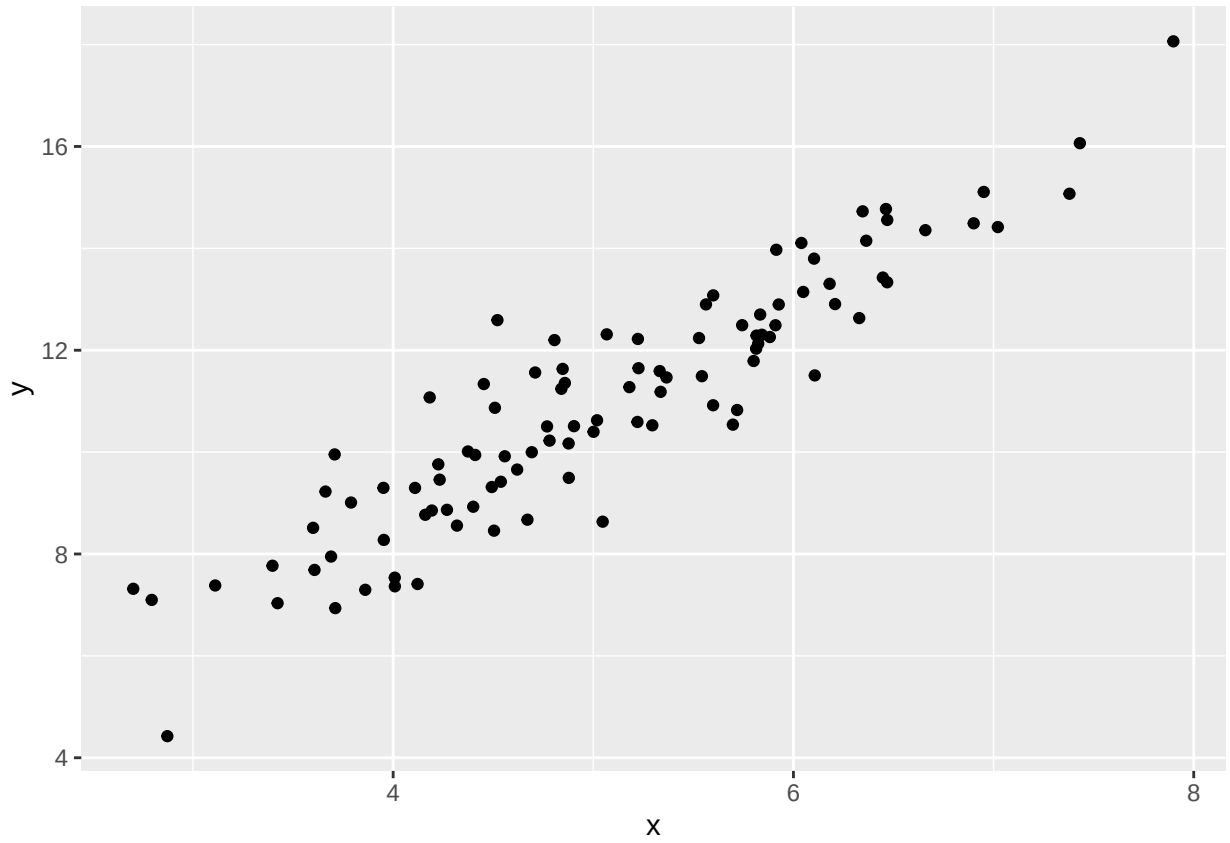
```
ggplot(data, aes(x)) + geom_histogram()
```

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



This is the scatter plot of the raw data

```
ggplot(data, aes(x,y)) + geom_point()
```



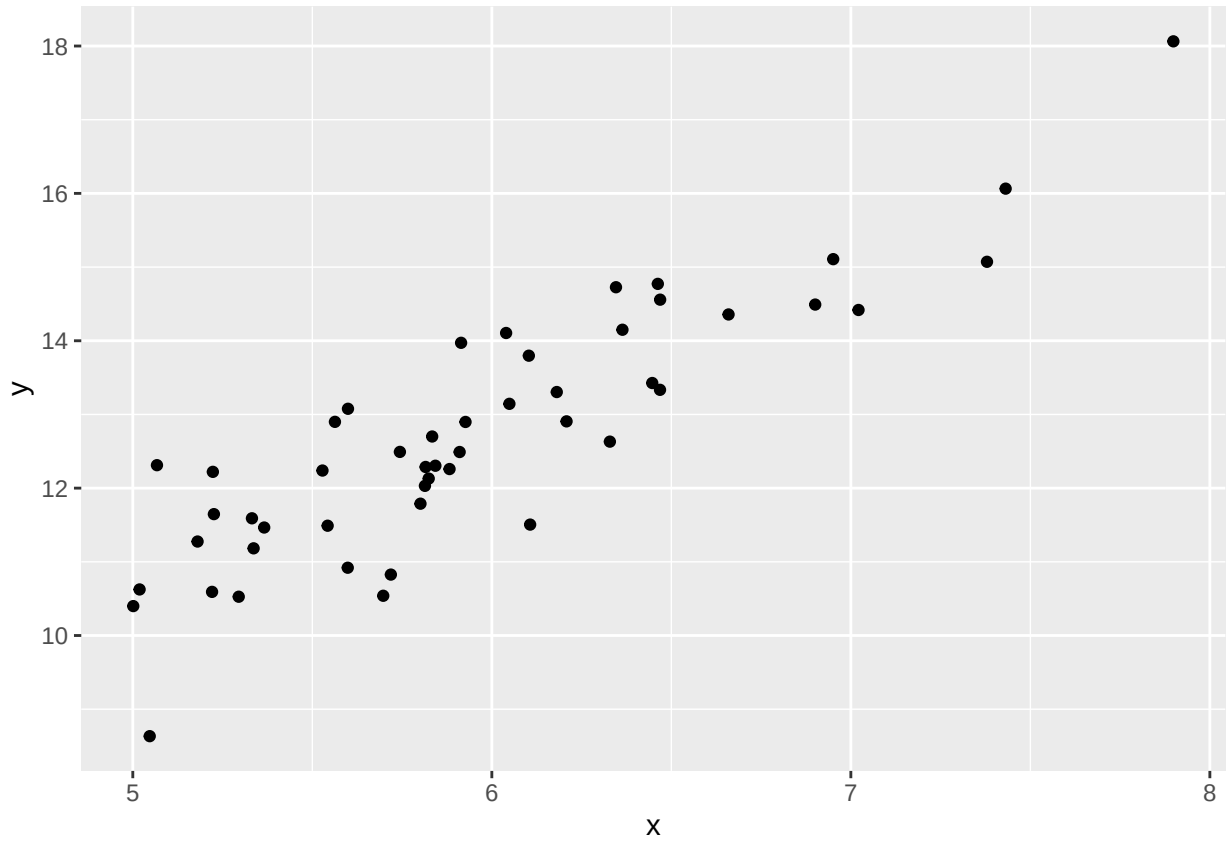
This is a filtered dataset for x > 5

```
filtered_data <- data %>% filter(x>5)

summary(filtered_data) # We can quickly find some statistics using the summary() function
```

```
##        x               y         
##  Min.   :5.001   Min.   : 8.634  
##  1st Qu.:5.532   1st Qu.:11.527  
##  Median :5.838   Median :12.491  
##  Mean   :5.954   Mean   :12.715  
##  3rd Qu.:6.342   3rd Qu.:13.929  
##  Max.   :7.898   Max.   :18.063
```

This is the scatter plot of the filtered data

```
ggplot(filtered_data, aes(x,y)) + geom_point()
```


18
16
14
12
10
y
5
6
7
8
x