Wecome to PSYCH 420

Your Name

2025-08-28

This is a heading

This is some descriptive text

```
# This is a comment
# I loaded some packages here. Now these packages are accessible to the rest of the document

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 becor
```

```
library(ggplot2)
```

If the PDF doesn't compile...

You may have to install "tinytex.' Un-comment the two lines below and run them with the green triangle on the right. Then re-comment them.

```
#install.packages('tinytex')
#tinytex::install_tinytex()
```

Here is some fake data

```
a <- 2 # We can use <- to assign values to variables

b <- 1

x <- rnorm(n = 100, mean = 5, sd=1) # Here, I sampled some random values from a normal distri

errors <- rnorm(n = 100, mean = 0, sd=1) # Same for the errors

y <- a*x+b + errors # I generated fake y values by using the function y = ax + b

data <- tibble(x = x, y = a*x+b + errors) # I created a dataframe to store the fake data
```

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

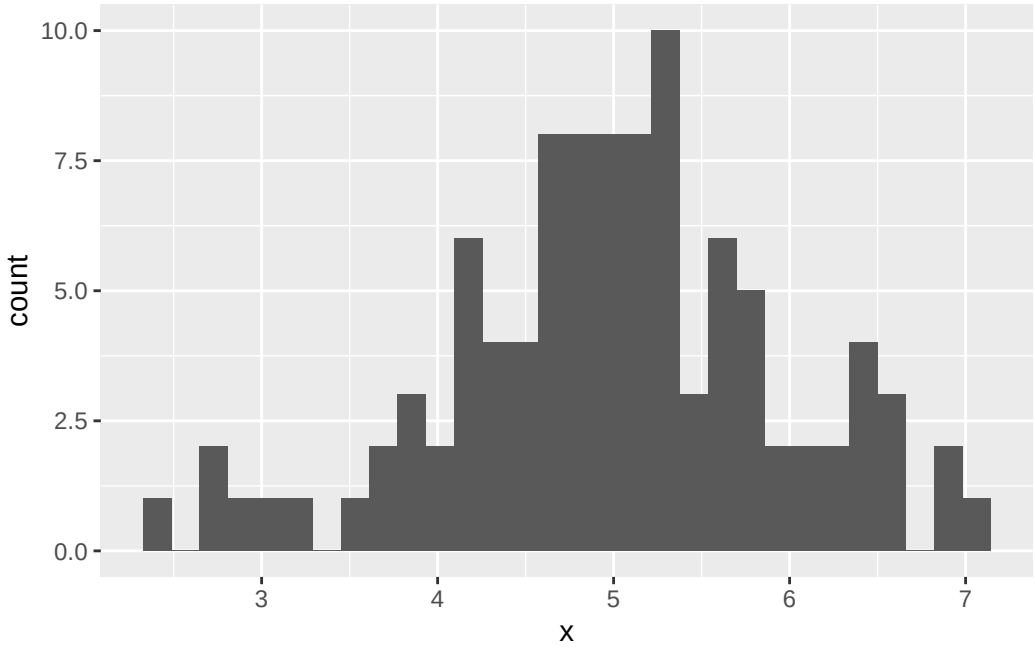
```
head(data)
```

```
# A tibble: 6 x 2
      x     y
  <dbl> <dbl>
1  5.22 11.5
2  5.26 10.6
3  5.07 10.6
4  6.89 15.5
5  4.35  9.68
6  5.56 12.3
```

This is the histogram of x

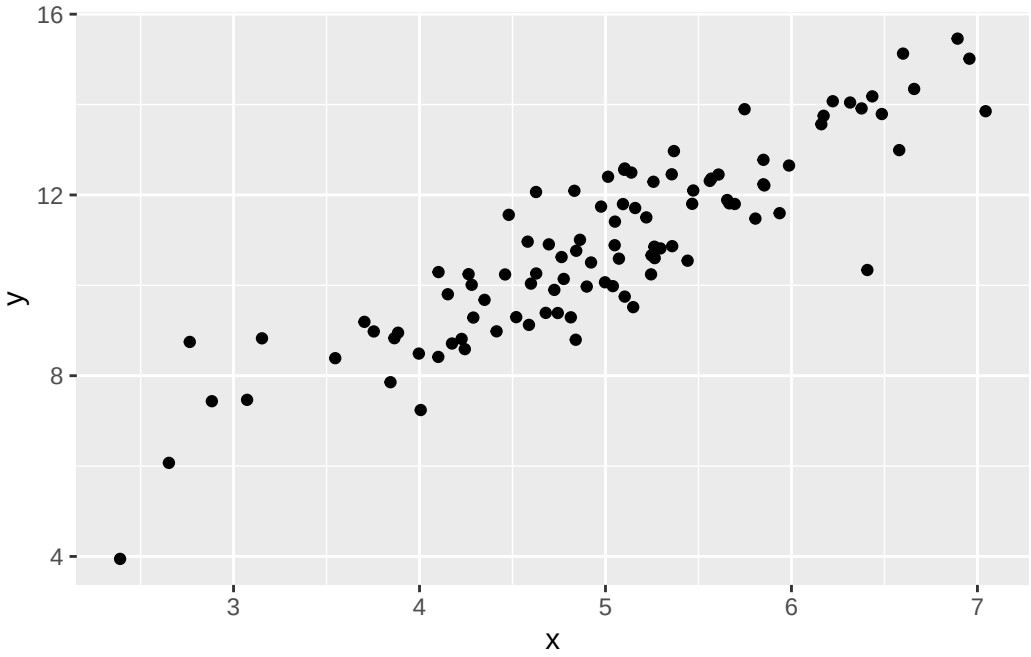
```
ggplot(data, aes(x)) + geom_histogram()
```

```
`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

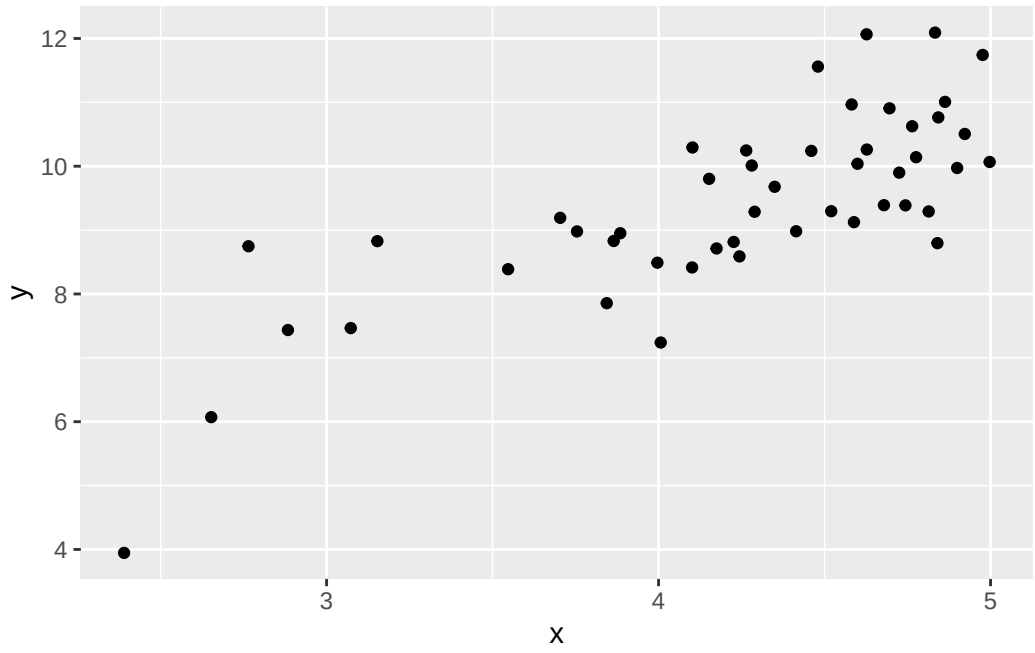
```
filtered_data <- data %>% filter(x<5)

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

```
       x               y         
 Min.   :2.390   Min.   : 3.946  
 1st Qu.:3.969   1st Qu.: 8.783  
 Median :4.382   Median : 9.341  
 Mean   :4.229   Mean   : 9.404  
 3rd Qu.:4.730   3rd Qu.:10.250  
 Max.   :4.997   Max.   :12.092  
```

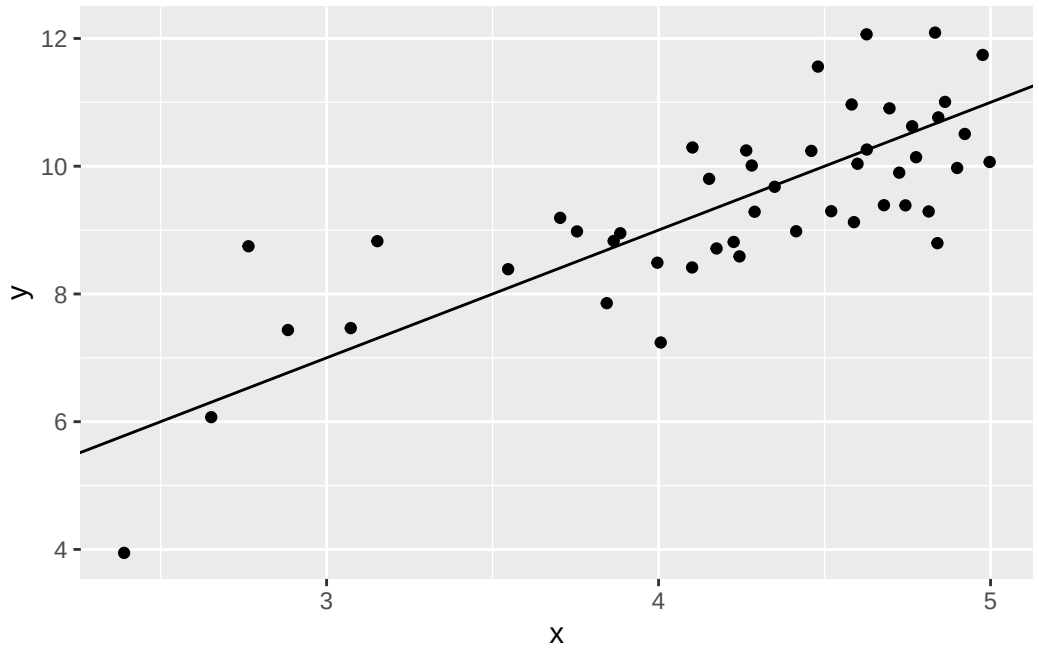
This is the scatter plot of the filtered data

```
filtered_scatter <- ggplot(filtered_data, aes(x,y)) + geom_point()
filtered_scatter
```



You can add a line of best fit like so

```
filtered_scatter + geom_abline(slope = a, intercept = b)
```


12
10
8
6
4
y
3
4
5
x