assignment 1

Installing packages and library

Load data - CSV

This is a COVID dataset and downloaded Open data

```
COVID19data <- readr::read_csv(file ="https://data.ontario.ca/dataset/8f3a449b-bde5-4631-ada6-8bd94dbc7
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   date = col_date(format = ""),
##   oh_region = col_character(),
##   ICU = col_double(),
##   ICU_vented = col_double(),
##   hospitalizations = col_double(),
##   icu_crci_total = col_double(),
##   icu_crci_total_vented = col_double(),
##   icu_former_covid = col_double(),
##   icu_former_covid_vented = col_double()
## )
```

head

```
head(COVID19data)
```

```
## # A tibble: 6 x 9
##   date       oh_region   ICU ICU_vented hospitalizations icu_crci_total
##   <date>     <chr>     <dbl>      <dbl>            <dbl>           <dbl>
## 1 2020-04-02 CENTRAL      51         39              113               0
## 2 2020-04-03 CENTRAL      57         52              141               0
## 3 2020-04-04 CENTRAL      65         57              143               0
## 4 2020-04-05 CENTRAL      65         55              174               0
## 5 2020-04-06 CENTRAL      73         58              187               0
## 6 2020-04-07 CENTRAL      83         68              183               0
## # ... with 3 more variables: icu_crci_total_vented <dbl>,
## #   icu_former_covid <dbl>, icu_former_covid_vented <dbl>
```

correlation

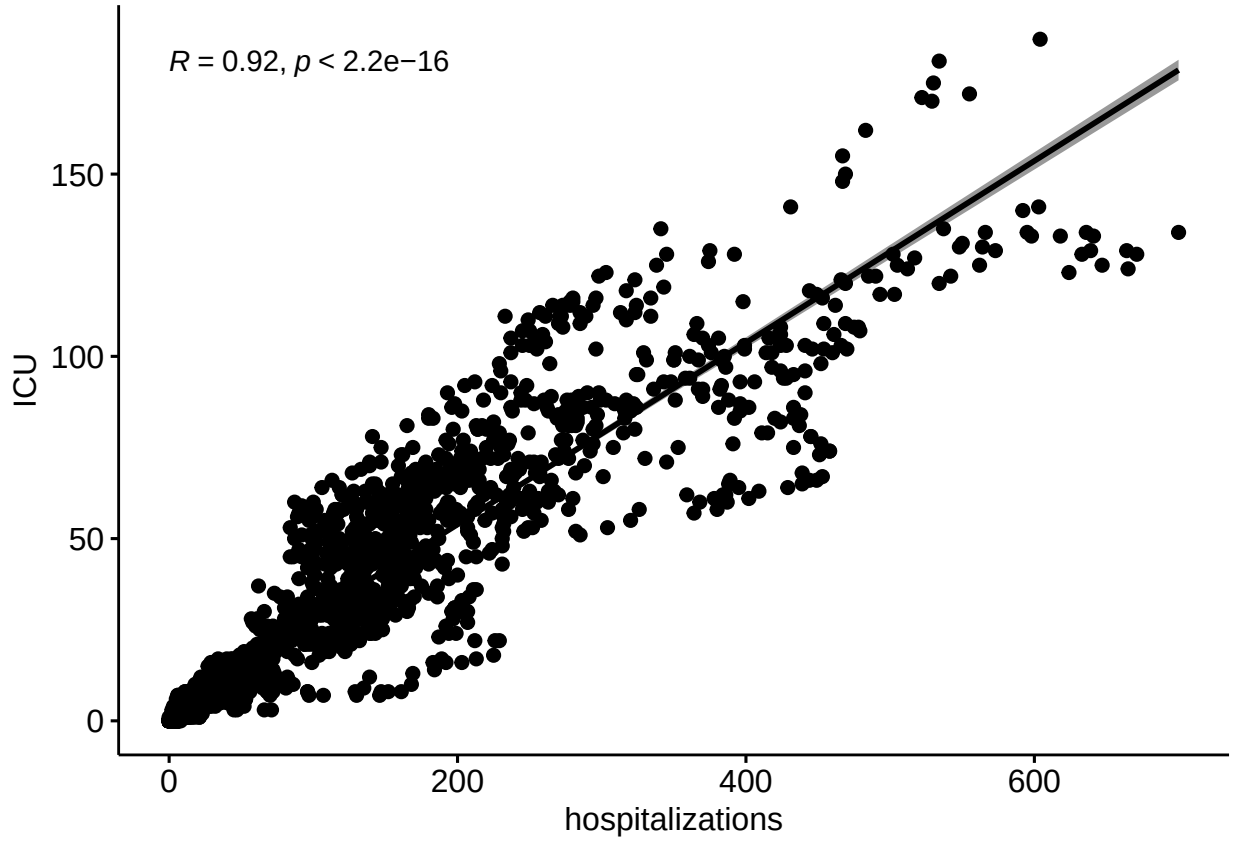
```
X = COVID19data[,"ICU_vented"]
Y = COVID19data[,"ICU"]
CORRELATION = cor(Y,X,method = "pearson")
CORRELATION
```

```
##     ICU_vented
## ICU  0.9841608
```

graph in correlation

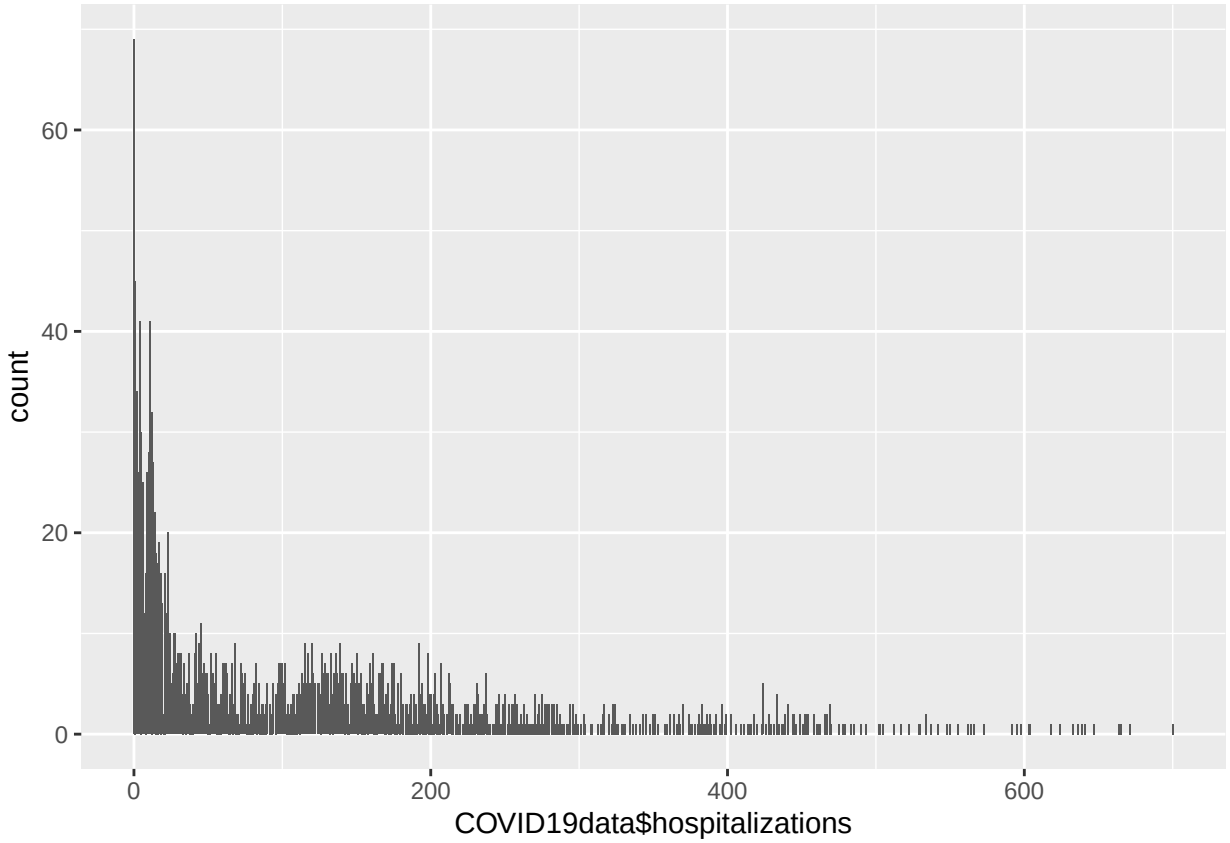
```
ggscatter(COVID19data, x = "hospitalizations", y = "ICU",add = "reg.line", conf.int = TRUE,cor.coef = TR
```

```
## `geom_smooth()` using formula 'y ~ x'
```



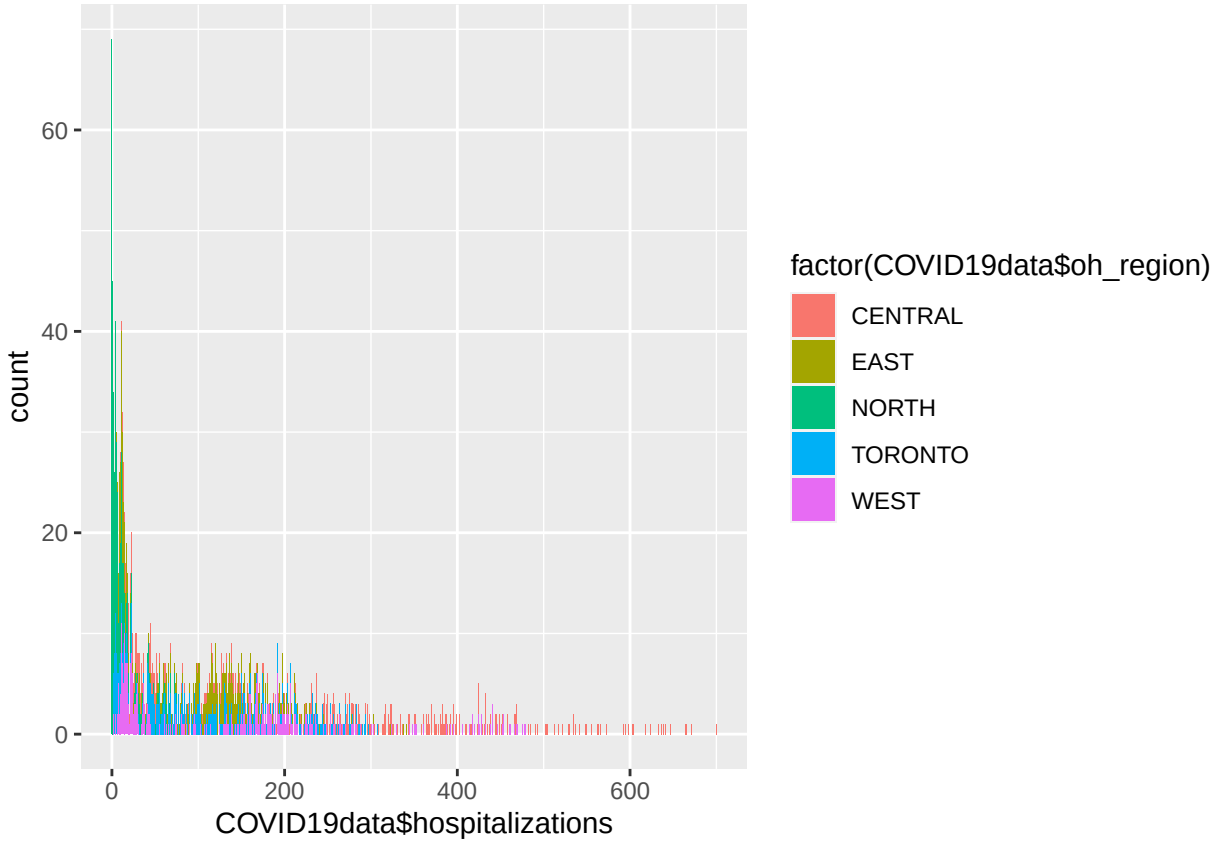
graph - bar plot hospitalisation

```
ggplot(data = COVID19data,aes(x = COVID19data$hospitalizations)) + geom_bar()
```



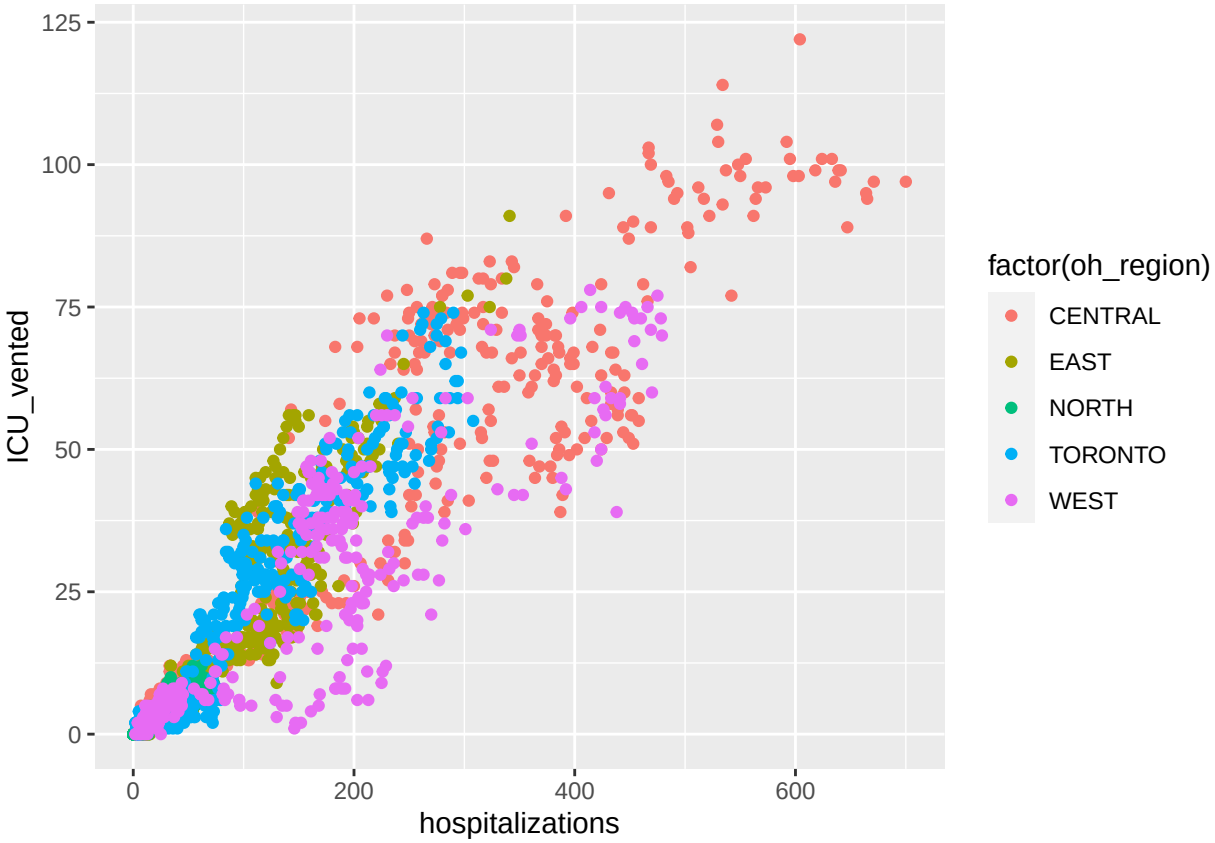
barplot with date

```
ggplot(data = COVID19data,aes(x = COVID19data$hospitalizations,fill = factor(COVID19data$oh_region))) +
```



scatter plot

```
ggplot(data = COVID19data,aes(x = hospitalizations,y = ICU_vented,col = factor(oh_region)))+ geom_point
```


125
100
75
50
25
0
ICU_vented
0
200
400
600
hospitalizations
factor(oh_region)
CENTRAL
EAST
NORTH
TORONTO
WEST