

Aufgaben Woche 1

Bearbeitet die folgenden Aufgaben:

- Speichere den Datensatz `airquality` in der Variable `airQuality`.
- Berechne die Gesamtdurchschnittstemperatur.
- Berechne die Durchschnittstemperatur für den Monat Juli.
- Vergleiche, ob die Monate Juli und Mai sich in ihrer Durchschnittstemperatur unterscheiden.

Hole Datensatz `airquality` und speicher den Data Frame in der Variable `airQuality`.

```
data("airquality")
airQuality <- airquality
head(airQuality)
```

```
##   Ozone Solar.R Wind Temp Month Day
## 1    41     190  7.4  67     5    1
## 2    36     118  8.0  72     5    2
## 3    12     149 12.6  74     5    3
## 4    18     313 11.5  62     5    4
## 5    NA      NA 14.3  56     5    5
## 6    28      NA 14.9  66     5    6
```

Berechne die Gesamtdurchschnittstemperatur.

```
#Ausgabe mit print und paste
airQuality_mean <- airQuality$Temp %>% mean()
print(paste0("Gesamtdurchschnittstemperatur: ", airQuality_mean))
```

```
## [1] "Gesamtdurchschnittstemperatur: 77.8823529411765"
```

```
#Ausgabe mit cat
airQuality_mean <- mean(airQuality$Temp)
cat('Gesamtdurchschnittstemperatur: ', airQuality_mean)
```

```
## Gesamtdurchschnittstemperatur: 77.88235
```

Lösung mit `summary`

```
# select(airQuality, Temp, Month) %>% filter(Month == 5) %>% summary(Temp)
summary(select((filter(select(airQuality, Temp, Month), Month == 5)), Temp))
```

```
##      Temp
##  Min.   :56.00
## 1st Qu.:60.00
##  Median :66.00
##   Mean  :65.55
## 3rd Qu.:69.00
##   Max.   :81.00
```

```
summary(select((filter(select(airQuality, Temp, Month), Month == 7)), Temp))
```

```
##      Temp
##  Min.   :73.0
## 1st Qu.:81.5
##  Median :84.0
##   Mean  :83.9
## 3rd Qu.:86.0
##   Max.   :92.0
```

Lösung mit summarise

```
# group_by(airquality, Month) %>% summarise(mean(Temp))
summarise(group_by(airQuality, Month), mean(Temp))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 5 x 2
##   Month `mean(Temp)`
##   <int>     <dbl>
## 1     5         65.5
## 2     6         79.1
## 3     7         83.9
## 4     8         84.0
## 5     9         76.9
```

Berechne die Durchschnittstemperatur für den Monat Mai.

```
#
airQuality_mean_5 <- airQuality %>% filter(Month == 5)
airQuality_mean_5 <- mean(airQuality_mean_5$Temp)
cat('Durchschnittstemperatur Mai: ', airQuality_mean_5)
```

```
## Durchschnittstemperatur Mai: 65.54839
```

Berechne die Durchschnittstemperatur für den Monat Juli.

```
#
airQuality_mean_7 <- airQuality %>% filter(Month == 7)
airQuality_mean_7 <- mean(airQuality_mean_7$Temp)
cat('Durchschnittstemperatur Juli: ', airQuality_mean_7)
```

```
## Durchschnittstemperatur Juli: 83.90323
```

Vergleiche, ob die Monate Juli und Mai sich in ihrer Durchschnittstemperatur unterscheiden.

```
gleich <- mean_diff <- (mean((airQuality %>% filter(Month == 7))$Temp)) == (
  mean((airQuality %>% filter(Month == 5))$Temp))
cat('gleich = ', gleich)
```

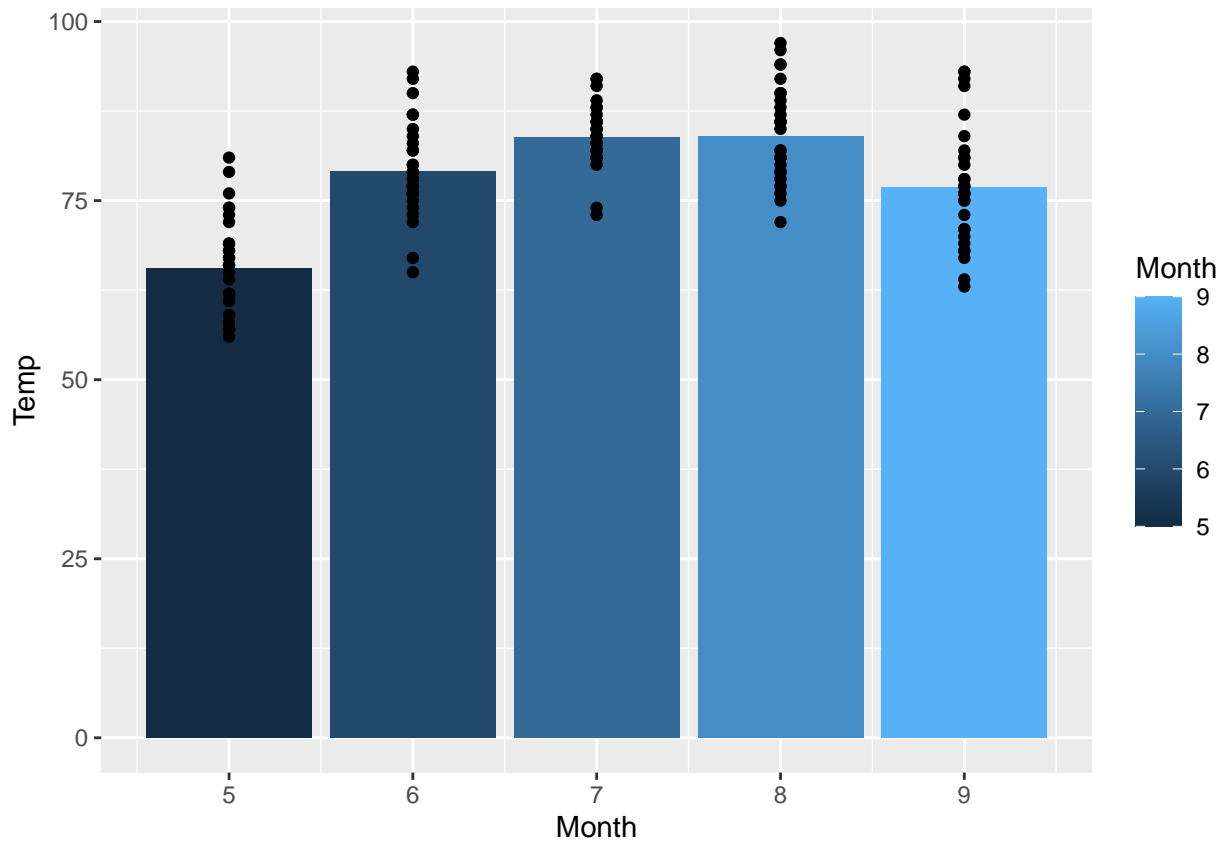
```
## gleich = FALSE
```

```
mean_diff <- (mean((airQuality %>% filter(Month == 7))$Temp)) - (
  mean((airQuality %>% filter(Month == 5))$Temp))
cat('Differenz in der Durchschnittstemperatur Juli gegenüber Mai: ', mean_diff)
```

```
## Differenz in der Durchschnittstemperatur Juli gegenüber Mai: 18.35484
```

Grafische Darstellung der Durchschnittstemperaturen

```
ggplot(airQuality, aes(Month, Temp, fill = Month)) +
  geom_bar(stat = 'summary', fun = "mean") +
  geom_point()
```



Bevorzugte Lösung

Lösung 2 ----

```
month <- c(unique(airQuality$Month))
mean_temp <- c()
for (mon in month) {
  mean_temp <- append(mean_temp, mean((airQuality %>% filter(Month == mon))$Temp))
  next
}
airQuality_mean_per_month <- data.frame(month, mean_temp)
airQuality_mean_per_month
```

```
##   month mean_temp
## 1     5  65.54839
## 2     6  79.10000
## 3     7  83.90323
## 4     8  83.96774
## 5     9  76.90000
```

#ggplot(airQuality_mean_per_month, aes(month, mean_temp)) + geom_point()

```
ggplot(airQuality_mean_per_month, aes(month, mean_temp)) +
  geom_bar(stat = 'summary', fun = 'mean', fill = 'lightblue', col = 'black') +
  geom_text(aes(label=mean_temp), vjust=-1, size=3) +
  theme(panel.border = element_blank(),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
```

```
panel.background = element_rect(fill = 'white')) +  
geom_line(stat = 'summary', fun = 'mean', col = "red", size = 2)
```

