Chapter 3 - Data Transformation with dplyr

Thalles Quinaglia Liduares

20/02/2022

Exercise 3.3 (incomplete)

Upload packages

```
library(dplyr)
library(nycflights13)
```

Upload database

```
data<-nycflights13::flights

attach(data)
```

3. How many flights have a missing dep_time? What other variables are missing? What might these rows represent?

```
miss_deptime<- data %>%
            filter(!is.na(dep_time)) %>%
            summarise(n=n())

miss_deptime
```

```
## # A tibble: 1 x 1
##        n
##    <int>
## 1 328521
```

```
length(dep_time)
```

```
## [1] 336776
```

```
number_of_na<-length(dep_time)-miss_deptime

number_of_na
```

```
##      n
## 1 8255
```

Hence, there's 8255 NA's in the `dep_time` variable.