Tutorial 3 DSA2101

Waseem

11/10/2021

```
library(jsonlite)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter()  masks stats::filter()
## x purrr::flatten() masks jsonlite::flatten()
## x dplyr::lag()     masks stats::lag()
```

```
#qn 1
```

1. Ensure that the date column is of class Date, and then add a column named num_arrivals with the number of arrivals per day.

```
er_json <- fromJSON("data/er_arrivals.json")
num_arrivals <- vapply(er_json$times,length,2L)
er_json$date<-as.Date(er_json$date,format="%Y-%m-%d")
er_json<-cbind(er_json,num_arrivals)

full_seq <- seq(as.Date("1963-02-01"),as.Date("1964-03-31"),by="1 day")
date<-full_seq[which(!(full_seq%in%er_json$date))]
times<-rep("NA",times=231)
num_arrivals<-rep(0,times=231)
empty_df<-data.frame(date,times,num_arrivals)
new_er_json <-rbind(er_json,empty_df)
new_er_json<-arrange(new_er_json, date)
```

2. Add rows to your data, containing the missing days, with NA's in the times column, and 0's in the num_arrivals column.

```
#qn 2
N<-425
Yk<-new_er_json$num_arrivals
Xk<-table(Yk)
Xk<-as.vector(Xk)
k<-0:4
```

3. Use the ER data to create the Poisson-ness plot for this dataset.

```r
phi<- lfactorial(k) + log(Xk/N)
plot(x=k,y=phi)
lm_out<-(lm(phi~k))
coefs<-coef(lm_out)
abline(a=coefs[1],b=coefs[2],lty=2)
```