PSET 02 - NBA Games

S&DS 361

Due 2024-02-06

Visualizing the NBA schedule

Let's visualize how often teams play each other in a season, to better understand the structure of the NBA schedule.

```
d = readRDS('data/games.rds')
d = d %>%
  filter(lg=='nba', season %in% 2022, season.type=='reg') %>%
  select(date, away, home, ascore, hscore, season, gid)
head(d)
```

```
##         date away home ascore hscore season      gid
## 1 2021-10-19  BKN  MIL    104    127   2022 22100001
## 2 2021-10-19  GSW  LAL    121    114   2022 22100002
## 3 2021-10-20  OKC  UTA     86    107   2022 22100011
## 4 2021-10-20  SAC  POR    124    121   2022 22100013
## 5 2021-10-20  DEN  PHX    110     98   2022 22100012
## 6 2021-10-20  ORL  SAS     97    123   2022 22100010
```

```
dg = d %>%
  group_by(away, home) %>%
  summarise(games = n()) %>%
  ungroup() %>%
  complete(away, home, fill=list(games=0)) ## new function!
```

```
## `summarise()` has grouped output by 'away'. You can override using the
## `.groups` argument.
```

```
head(dg)
```

```
## # A tibble: 6 x 3
##   away  home  games
##   <chr> <chr> <int>
## 1 ATL   ATL       0
## 2 ATL   BKN       1
## 3 ATL   BOS       2
## 4 ATL   CHA       2
## 5 ATL   CHI       2
## 6 ATL   CLE       2
```

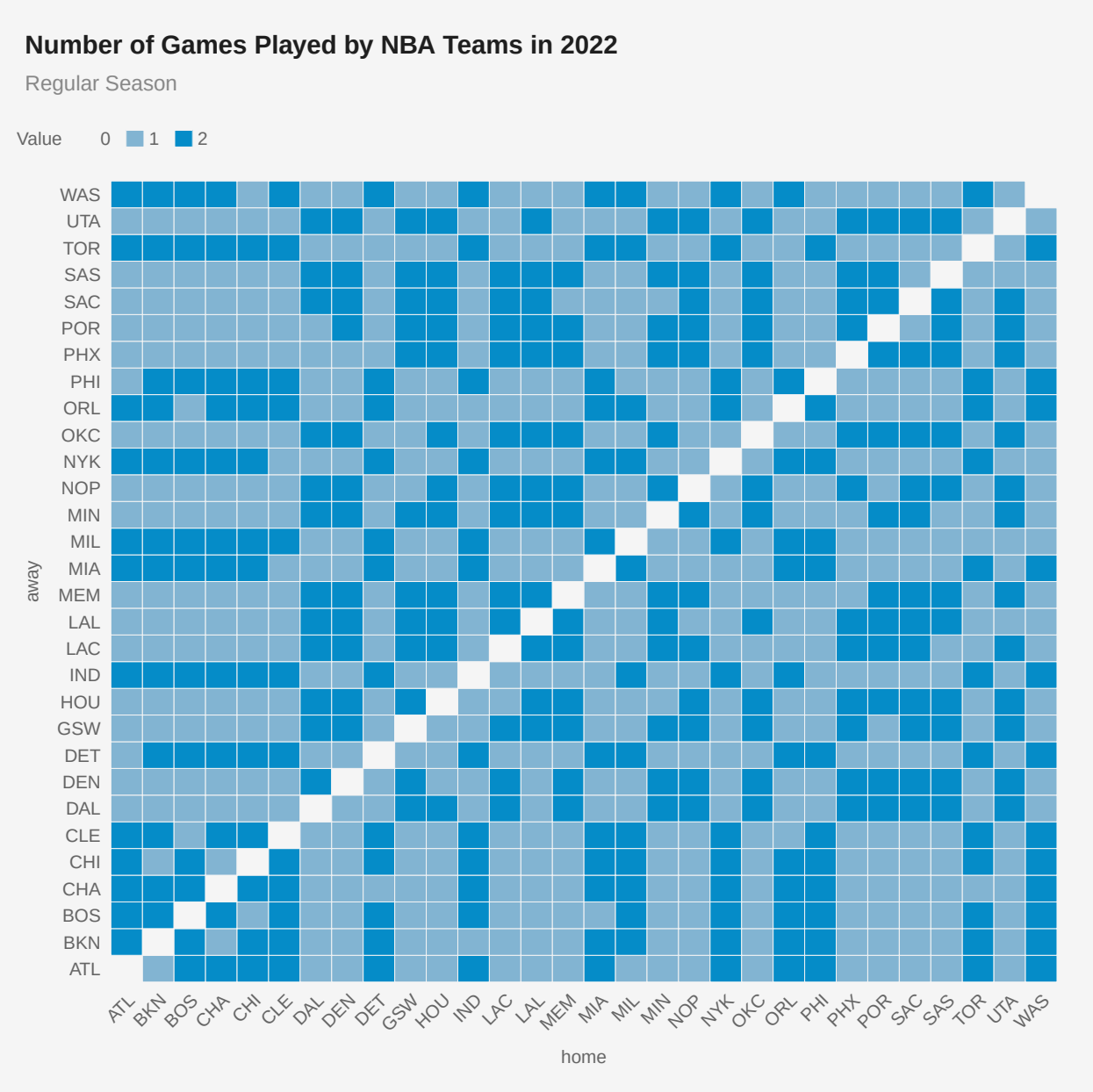
Visualizing the schedule with a grid plot

```
title = "Number of Games Played by NBA Teams in 2022"

gg = ggplot(dg, aes(x = home,
                    y = away,
                    fill = as.character(games))) +
  geom_tile(show.legend = TRUE) +
  scale_fill_manual(values = c(pubbackgray, publightblue, pubblue))+
  geom_tile(color = pubbackgray) +
  labs(title    = title,
       subtitle = "Regular Season",
       fill = 'Value')

gg %>% pub(type = 'grid') +
  scale_x_discrete(
    expand = expansion(mult = c(0,0)), guide = guide_axis(angle = 45),
    position = "bottom"
  )
```

```
## Scale for x is already present.
## Adding another scale for x, which will replace the existing scale.
```


Number of Games Played by NBA Teams in 2022
Regular Season
Value 0 1 2
away
WAS
UTA
TOR
SAS
SAC
POR
PHX
PHI
ORL
OKC
NYK
NOP
MIN
MIL
MIA
MEM
LAL
LAC
IND
HOU
GSW
DET
DEN
DAL
CLE
CHI
CHA
BOS
BKN
ATL
ATL BKN BOS CHA CHI CLE DAL DEN DET GSW HOU IND LAC LAL MEM MIA MIL MIN NOP NYK OKC ORL PHI PHX POR SAC SAS TOR UTA WAS
home

Ordering the teams by division

```
tms = read.csv('data/nba.teams.csv')

tms = tms %>%
  arrange(conf, div) %>%
  mutate(conf = paste0(toupper(substr(conf, 1, 1)),
                               substr(conf, 2, nchar(conf))),

         div  = paste0(toupper(substr(div , 1, 1)),
                               substr(div , 2, nchar( div))),

         div  = factor(div,
                       levels = unique(div)))
head(tms)
```

```
##   team      div conf
## 1  BKN Atlantic East
## 2  BOS Atlantic East
## 3  NYK Atlantic East
## 4  PHI Atlantic East
## 5  TOR Atlantic East
## 6  CHI  Central East
```

```
dg2 = dg %>%
  left_join(select(tms, team, div), by=c('home' = 'team')) %>%
  rename(home_div = div) %>%
  left_join(select(tms, team, div), by=c('away' = 'team')) %>%
  rename(away_div = div)

teams.order = tms %>%
  select(team) %>%
  unlist()
head(teams.order)
```

```
## team1 team2 team3 team4 team5 team6
## "BKN" "BOS" "NYK" "PHI" "TOR" "CHI"
```

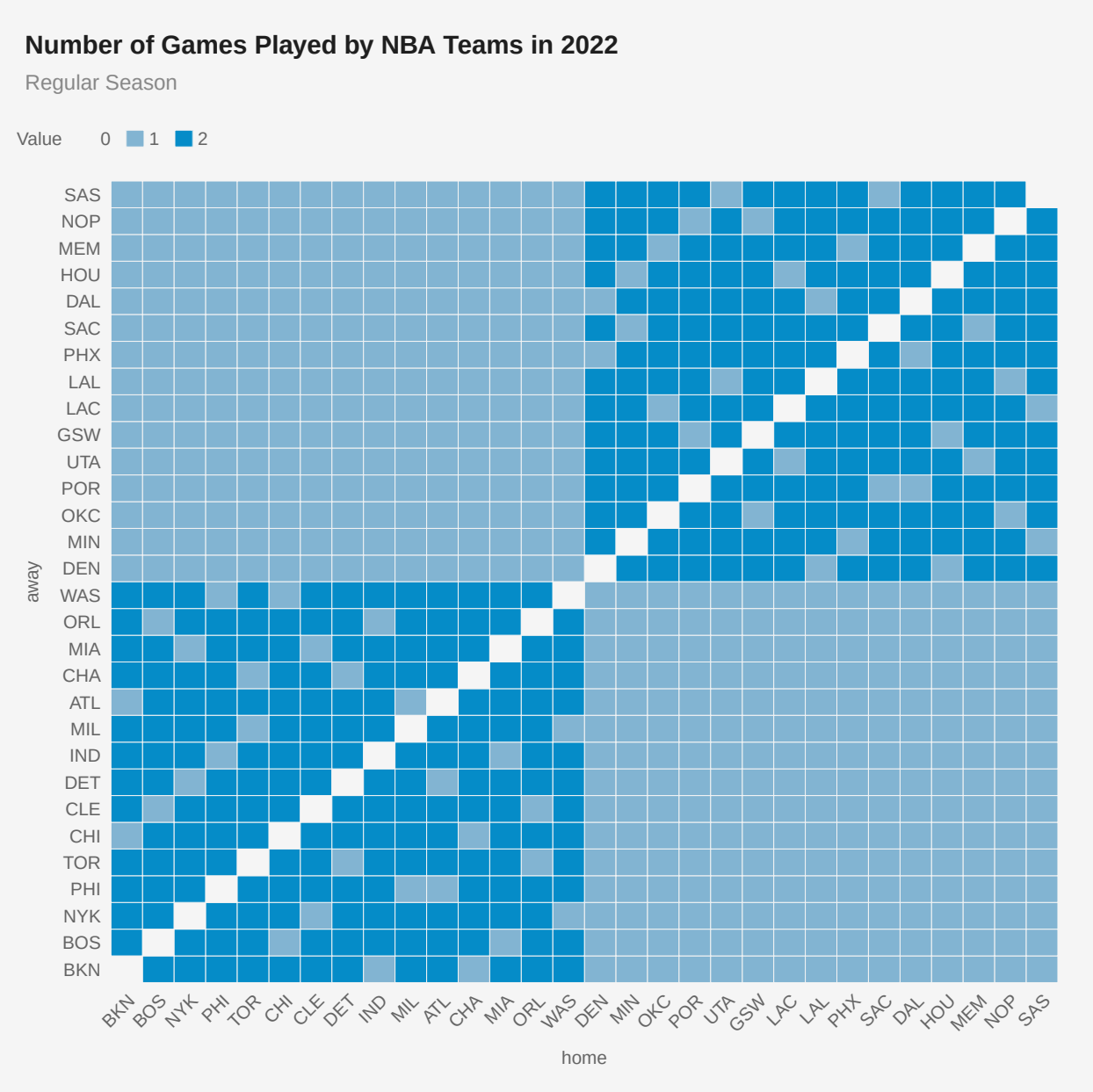
```
dg2 = dg2 %>%
  mutate(home = factor(home, levels = teams.order),
         away = factor(away, levels = teams.order))

gg2 = ggplot(dg2, aes(x = home,
                y = away,
                fill = as.character(games))) +
  geom_tile(show.legend = TRUE) +
  scale_fill_manual(values = c(pubbackgray, publightblue, pubblue))+
  geom_tile(color = pubbackgray) +
  labs(title    = title,
       subtitle = "Regular Season",
       fill = 'Value')

gg2 %>% pub(type = 'grid') +
  scale_x_discrete(
    expand = expansion(mult = c(0,0)), guide = guide_axis(angle = 45),
```

```
    position = "bottom"
  )
```

```
## Scale for x is already present.
## Adding another scale for x, which will replace the existing scale.
```

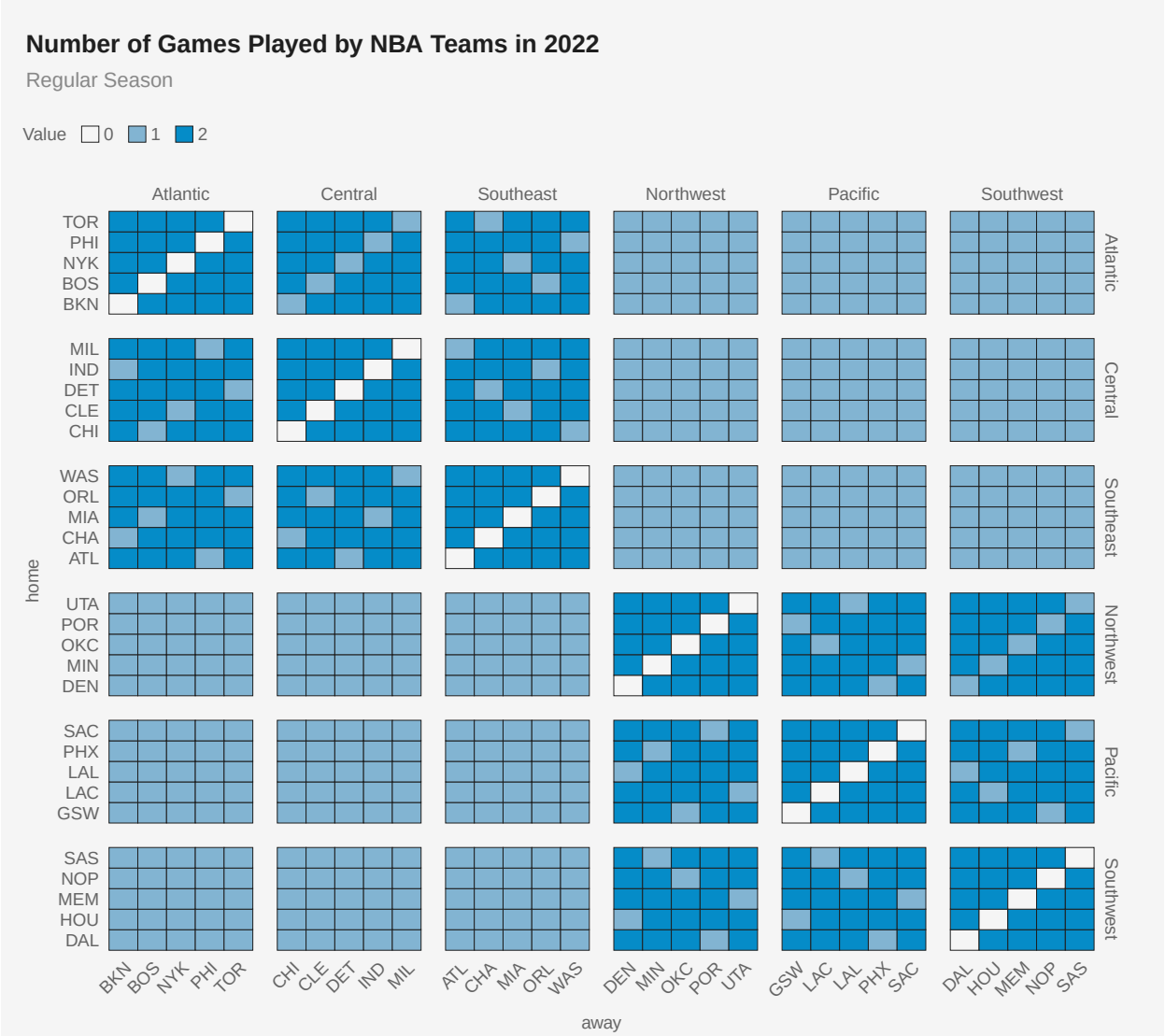


Creating separation between the divisions

```
g3 = ggplot(dg2,
            aes(x = away,
                y = home,
                fill = as.character(games))) +
  geom_tile(show.legend = TRUE, color = pubdarkgray) +
  facet_grid(home_div ~ away_div, scales = 'free', space = 'free') +
  scale_fill_manual(values = c(pubbackgray, publightblue, pubblue))+
  labs(title    = title,
       subtitle = "Regular Season",
       fill = 'Value')

g3 %>% pub(type = 'grid') +
  scale_x_discrete(
    expand = expansion(mult = c(0,0)), guide = guide_axis(angle = 45),
    position = "bottom"
  )
```

```
## Scale for x is already present.
## Adding another scale for x, which will replace the existing scale.
```


Number of Games Played by NBA Teams in 2022
Regular Season
Value 0 1 2
Atlantic
Central
Southeast
Northwest
Pacific
Southwest
TOR PHI NYK BOS BKN
MIL IND DET CLE CHI
WAS ORL MIA CHA ATL
UTA POR OKC MIN DEN
SAC PHX LAL LAC GSW
SAS NOP MEM HOU DAL
home
BKN BOS NYK PHI TOR
CHI CLE DET IND MIL
ATL CHA MIA ORL WAS
DEN MIN OKC POR UTA
GSW LAC LAL PHX SAC
DAL HOU MEM NOP SAS
away