

Proyecto Final

Bloque 2: Estadística y Programación en R

Equipo 2

TEAM MEMBERS



Fernando Guevara



Gil Rodríguez



Ana Rosado



Karla Gómez



José Muñoz

Postwork 01: Probabilidad Marginal

```
library(dplyr)
Aa<-read.csv("https://www.football-data.co.uk/mmz4281/1920/SP1.csv")
GolesLocal<-data.frame(Aa$FTHG); GolesVisitante<-data.frame(Aa$FTAG)
P_Local_G=data.frame(table(GolesLocal)/JJG); P_Visit_G=data.frame(table(GolesVisitante)/JJG)
```

Goles	Probabilidad
0	0.2316
1	0.3474
2	0.2605
3	0.1000
4	0.0368
5	0.0211
6	0.0026

Goles	Probabilidad
0	0.3579
1	0.3526
2	0.2132
3	0.0474
4	0.0237
5	0.0053

Postwork 01: Probabilidad Conjunta

```
GolesLocalVisit <- table(GolesLocalVisit)
```

```
GolesLocalVisit <- GolesLocalVisit/JG
```

		Goles de Visitante					
		0	1	2	3	4	5
G.Local	0	0.0868	0.0737	0.0395	0.0211	0.0053	0.0053
	1	0.1132	0.1289	0.0842	0.0132	0.0079	0.0000
	2	0.1026	0.0921	0.0526	0.0079	0.0053	0.0000
	3	0.0368	0.0368	0.0184	0.0053	0.0026	0.0000
	4	0.0105	0.0132	0.0105	0.0000	0.0026	0.0000
	5	0.0053	0.0079	0.0079	0.0000	0.0000	0.0000
	6	0.0026	0.0000	0.0000	0.0000	0.0000	0.0000

Postwork 02

```
NewTemUn<-lapply(TemporadasUnidas,function(x){select(x,c('Date',  
'HomeTeam', 'AwayTeam','FTHG','FTAG','FTR'))})
```

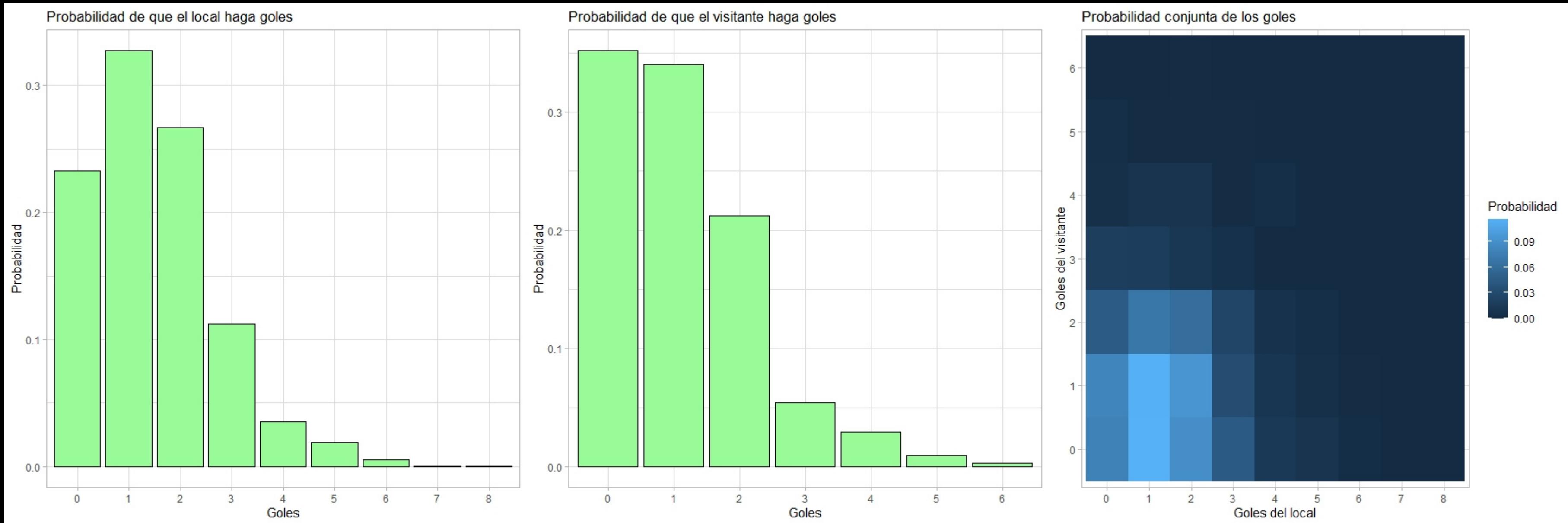
```
NewTemUn<-lapply(NewTemUn,function(x){mutate(x,Date=as.Date(Date, "%d/%m/%Y"))})
```

```
TemUn<-do.call(rbind,NewTemUn)
```

```
TemUn$Date<-as.Date(str_replace(TemUn$Date,'00','20'))
```

Fecha	Local	Visitante	G. Local	G. Visitante	Ganador
18/08/2017	Leganes	Alaves	1	0	H
18/08/2017	Valencia	Las Palmas	1	0	H
19/08/2017	Celta	Sociedad	2	3	A
19/08/2017	Girona	Ath Madrid	2	2	D
19/08/2017	Sevilla	Espanol	1	1	D
20/08/2017	Ath Bilbao	Getafe	0	0	D
...

Postwork 03



```
ggplot(ProbMargLocal, aes(x = G.Local, y = Probabilidad)) +
  geom_bar(stat="identity", col="black", fill="Pale green") +
  ggttitle("Probabilidad de que el local haga goles") +
  ylab("Probabilidad") +
  xlab("Goles") +
  theme_light()
```

```
ggplot(ProbMargVisit, aes(x = G.Visit, y = Probabilidad)) +
  geom_bar(stat="identity", col="black", fill="Pale green") +
  ggttitle("Probabilidad de que el visitante haga goles") +
  ylab("Probabilidad") +
  xlab("Goles") +
  theme_light()
```

```
ggplot(ProbConjunta, aes(x = G.Local, y = G.Visit, fill = Probabilidad)) +
  geom_tile() +
  ggttitle("Probabilidad conjunta de los goles") +
  labs(x="Goles del local",y="Goles del visitante",fill='Probabilidad') +
  theme_light()
```

Postwork 04

```
for (i in 1:100) {  
  Compuesto<-rnorm(nrow(Cocientes), mean=MediaCocientes, sd=SdCocientes)  
  Bootsmap<-cbind(Bootsmap,Compuesto)  
}  
Bootsmap<-apply(Bootsmap,2,sort)  
MediasBM<-apply(Bootsmap,2,mean)  
SDBM<-apply(Bootsmap,1,sd)  
  
z<-(1-Cocientes[,ncol(Cocientes)])/(SDBM/10)  
z0.025<- qnorm(p=0.025, lower.tail = FALSE)  
Resultz<- (z>-z0.025) & (z<z0.025)  
Marcadores<-Cocientes[Resultz,1:3]
```

G.Local	G.Visit	Cociente
0	0	0.9548
1	0	1.0061
0	1	1.0200
1	1	1.0319
2	1	1.0341
6	1	0.9794
1	2	0.9851
3	2	1.0305
1	3	0.9859
3	3	1.0055

Postwork 05

```
SmallData<data.frame(date=Date,home.team=HomeTeam,  
home.score=FTHG, away.team=AwayTeam,  
away.score=FTAG)
```

```
ranking <-rank.teams(scores=anotaciones,  
teams=equipos,  
max.date= fecha[n-1], min.date=fecha[1],  
date.format="%Y-%m-%d")
```

Team	Rankings	based	on matches	18/08/2017	to	16/07/2020	
						team	total
1	Barcelona	1.32	2.14	1.16	Barcelona	75	
2	Ath Madrid	1.07	1.3	1.62	Ath Madrid	75	
3	Real Madrid	1.06	1.63	1.28	Real Madrid	75	
4	Sevilla	0.66	1.44	1.1	Sevilla	75	
5	Getafe	0.52	1.14	1.26	Getafe	75	
6	Valencia	0.32	1.22	1.03	Valencia	75	

Postwork 05

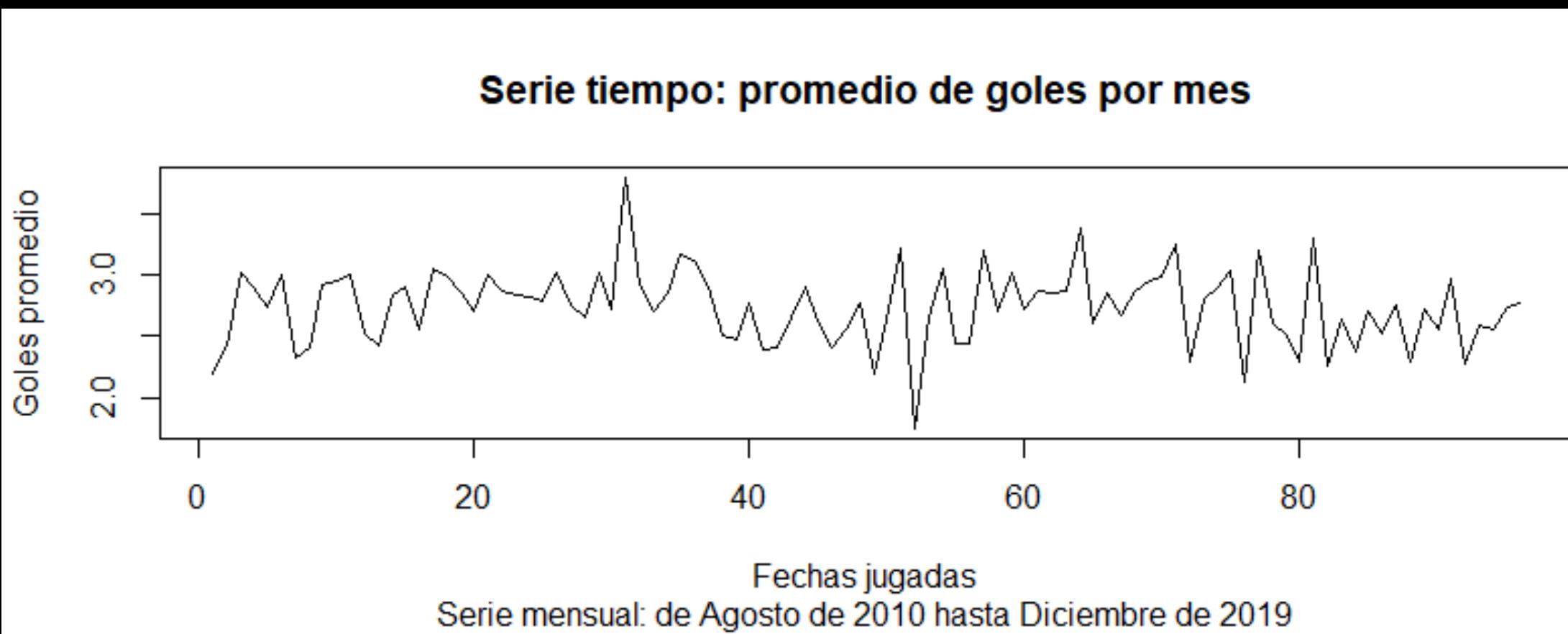
predict(ranking, max.date=fecha[n], min.date=fecha[n])

```
~/Curso BEDU/Curso_R/PW/postw5/ ~
Predicted Match Results for 2020-07-19 to 2020-07-19
Model based on data from 2017-08-18 to 2020-07-16
-----
2020-07-19 Alaves vs Barcelona, HW 9%, AW 76%, T 15%, pred score 0.7-2.5 actual: AW (0-5)
2020-07-19 Valladolid vs Betis, HW 29%, AW 43%, T 28%, pred score 1-1.3 actual: HW (2-0)
2020-07-19 Villarreal vs Eibar, HW 45%, AW 30%, T 26%, pred score 1.5-1.2 actual: HW (4-0)
2020-07-19 Ath Madrid vs Sociedad, HW 54%, AW 20%, T 27%, pred score 1.5-0.8 actual: T (1-1)
2020-07-19 Espanol vs Celta, HW 32%, AW 41%, T 27%, pred score 1.2-1.4 actual: T (0-0)
2020-07-19 Granada vs Ath Bilbao, HW 40%, AW 31%, T 29%, pred score 1.2-1 actual: HW (4-0)
2020-07-19 Leganes vs Real Madrid, HW 13%, AW 66%, T 21%, pred score 0.7-1.9 actual: T (2-2)
2020-07-19 Levante vs Getafe, HW 25%, AW 48%, T 27%, pred score 0.9-1.4 actual: HW (1-0)
2020-07-19 Osasuna vs Mallorca, HW 48%, AW 28%, T 25%, pred score 1.6-1.1 actual: T (2-2)
2020-07-19 Sevilla vs Valencia, HW 34%, AW 40%, T 26%, pred score 1.2-1.4 actual: HW (1-0)
> |
```

Postwork 06

```
Promedio<- df %>% group_by(Year.month) %>%  
summarise(gols.month=mean(sumagoles),  
Games.month=n())
```

Year.month	gols.month	Games.month
2010/08	2.2	10
2010/09	2.42	40
2010/10	3.03	39
2010/11	2.9	41
2010/12	2.73	30
2011/01	3	50



```
PromedioHastaDic2019<-  
ts(subset(Promedio, Year.month<='2019/12',  
select='gols.month'))
```

Postwork 07

```
con <- mongo(
```

```
  collection = "match",
  url = "mongodb://localhost"
)
```

```
con$insert(match_games)
```

```
con$count()
```

```
con$find(query='{"date": "2015-12-20",
```

```
          "$or": [{"home_team": "Real Madrid"}, {"away_team": "Real
Madrid"}]})
```

```
> con$find(query='{"date": "2015-12-20",
+               "$or": [{"home_team": "Real Madrid"}, {"away_team": "Real Madrid"}]}')
   date   home_team home_score away_team away_score
1 2015-12-20 Real Madrid           10 Vallecano            2
```

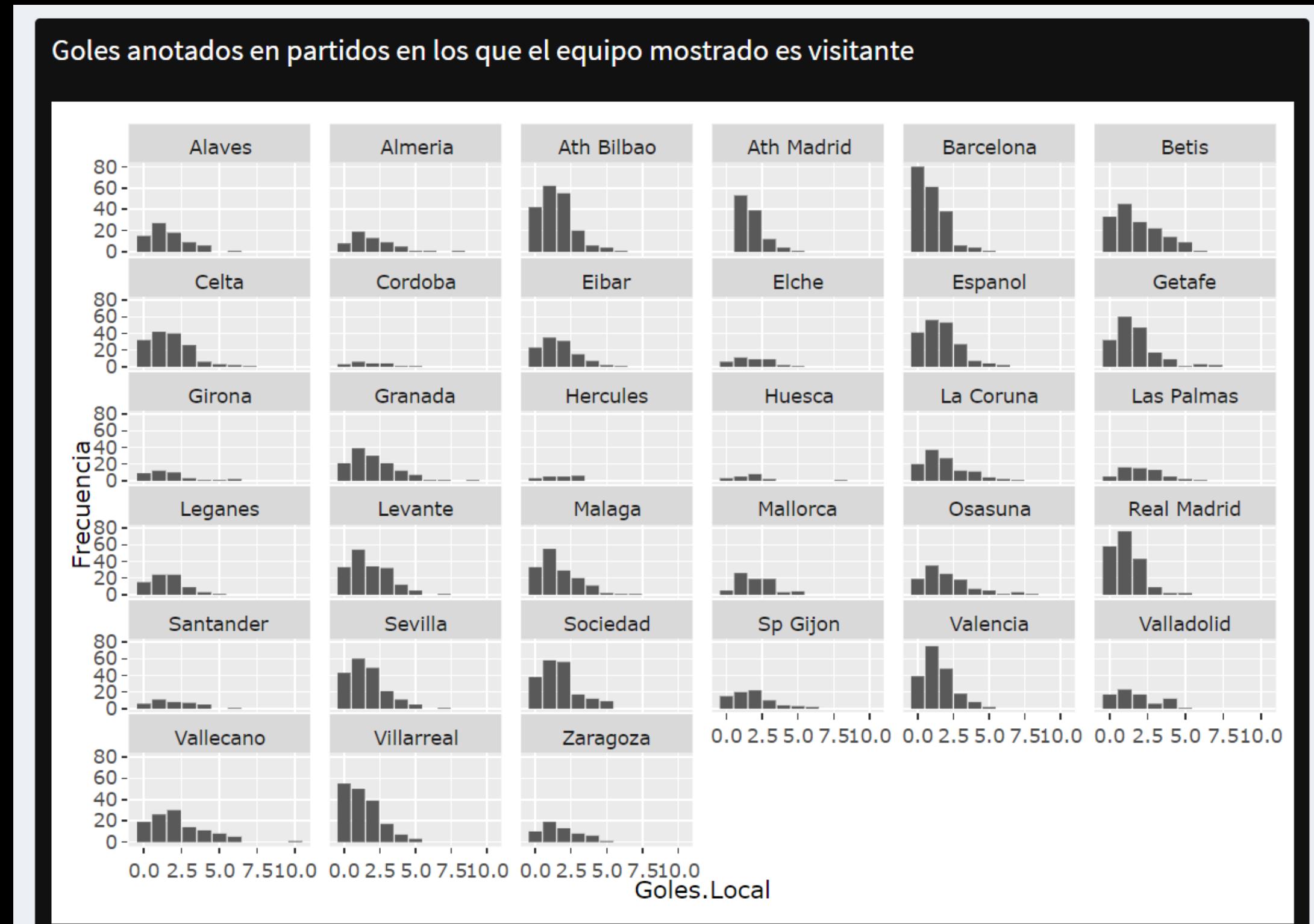
Postwork 08



SCAN ME

<https://gilesitorr.shinyapps.io/Postwork8/>

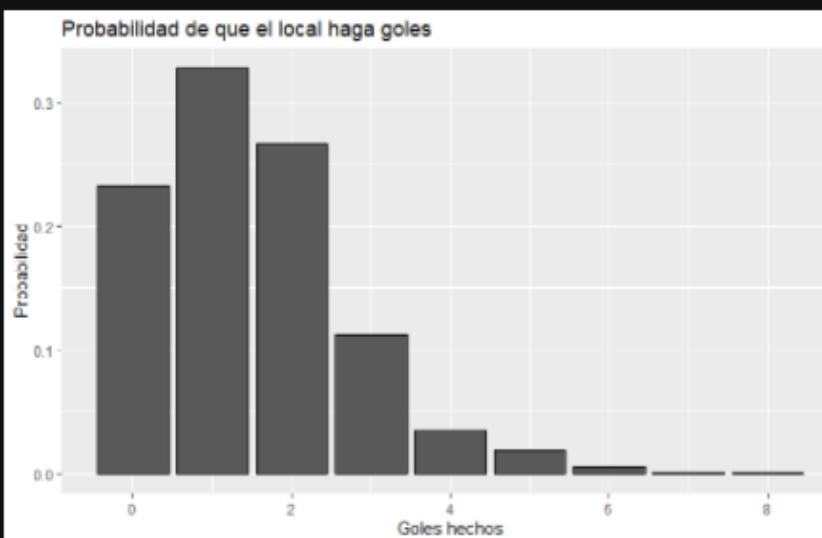
Postwork 08



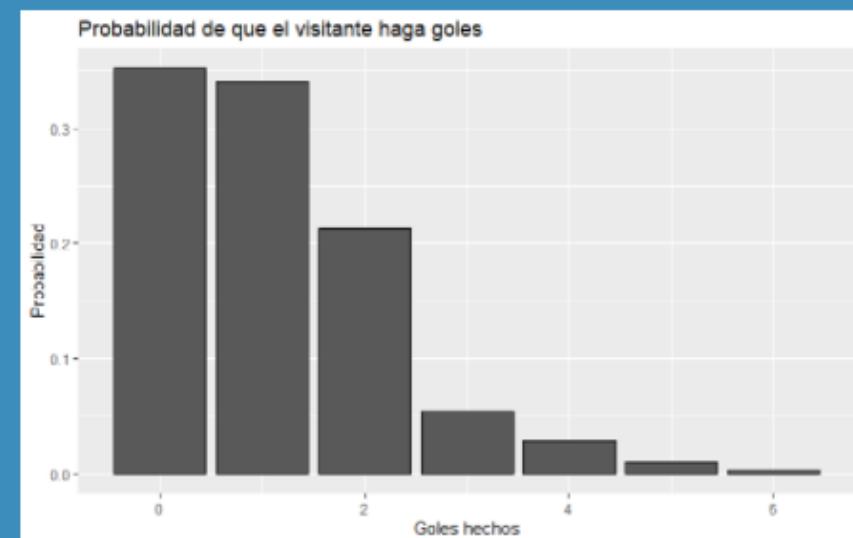
Postwork 08

Probabilidad de goles (2017-2020)

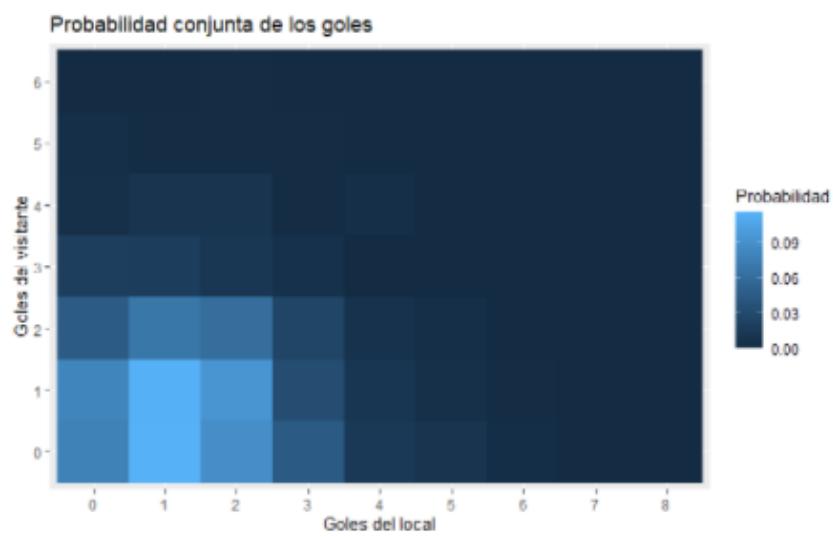
Probabilidad del local



Probabilidad del visitante



Probabilidad conjunta



Postwork 08

Marcadores de cada partido (2010-2020)

Show entries

Search:

Fecha	Equipo Local	Goles Local	Equipo Visitante	Goles Visitante
2010-08-28	Hercules	0	Ath Bilbao	1
2010-08-28	Levante	1	Sevilla	4
2010-08-28	Malaga	1	Valencia	3
2010-08-29	Espanol	3	Getafe	1
2010-08-29	La Coruna	0	Zaragoza	0

Fecha

Equipo Local

Goles Local

Equipo Visitante

Goles Visitante

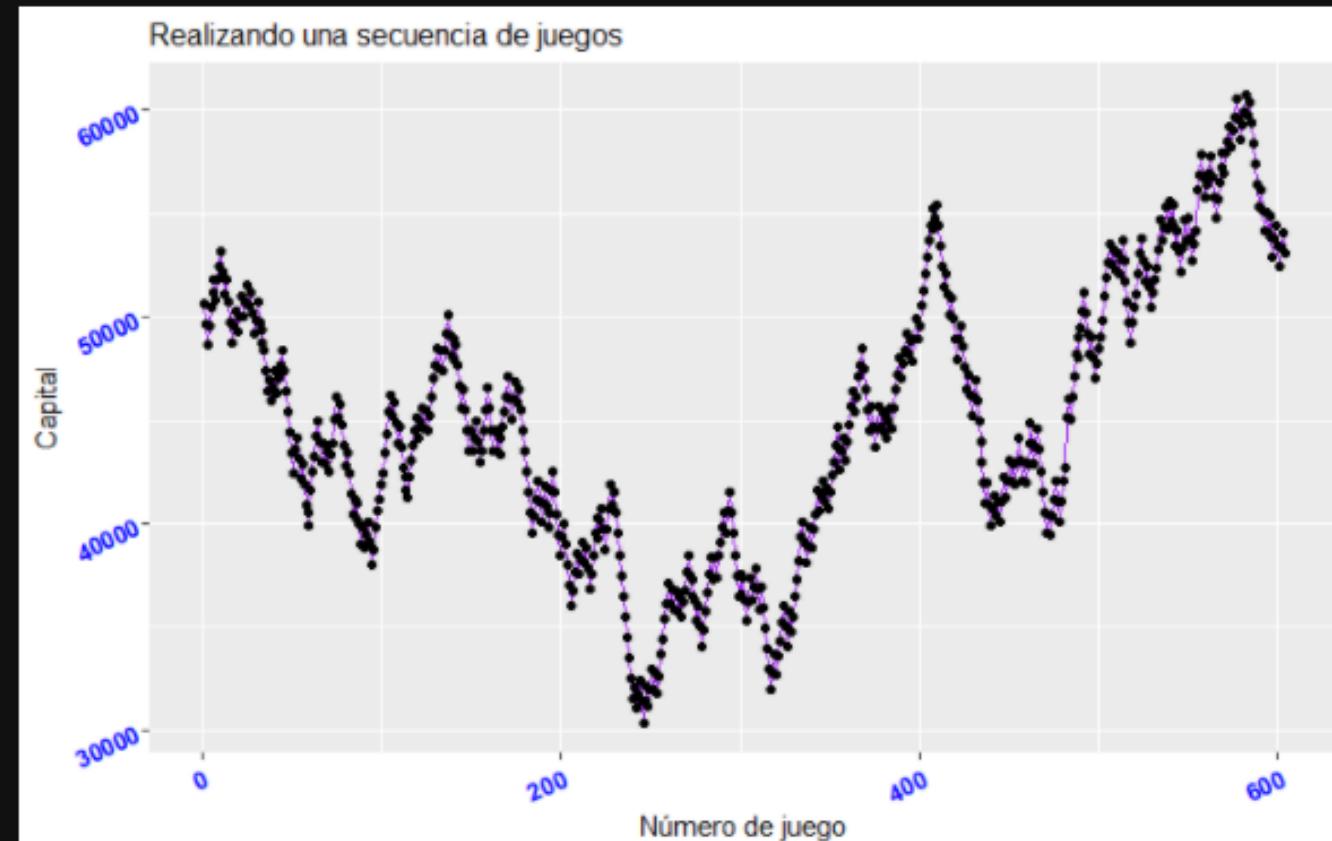
Showing 1 to 5 of 3,800 entries

Previous ... Next

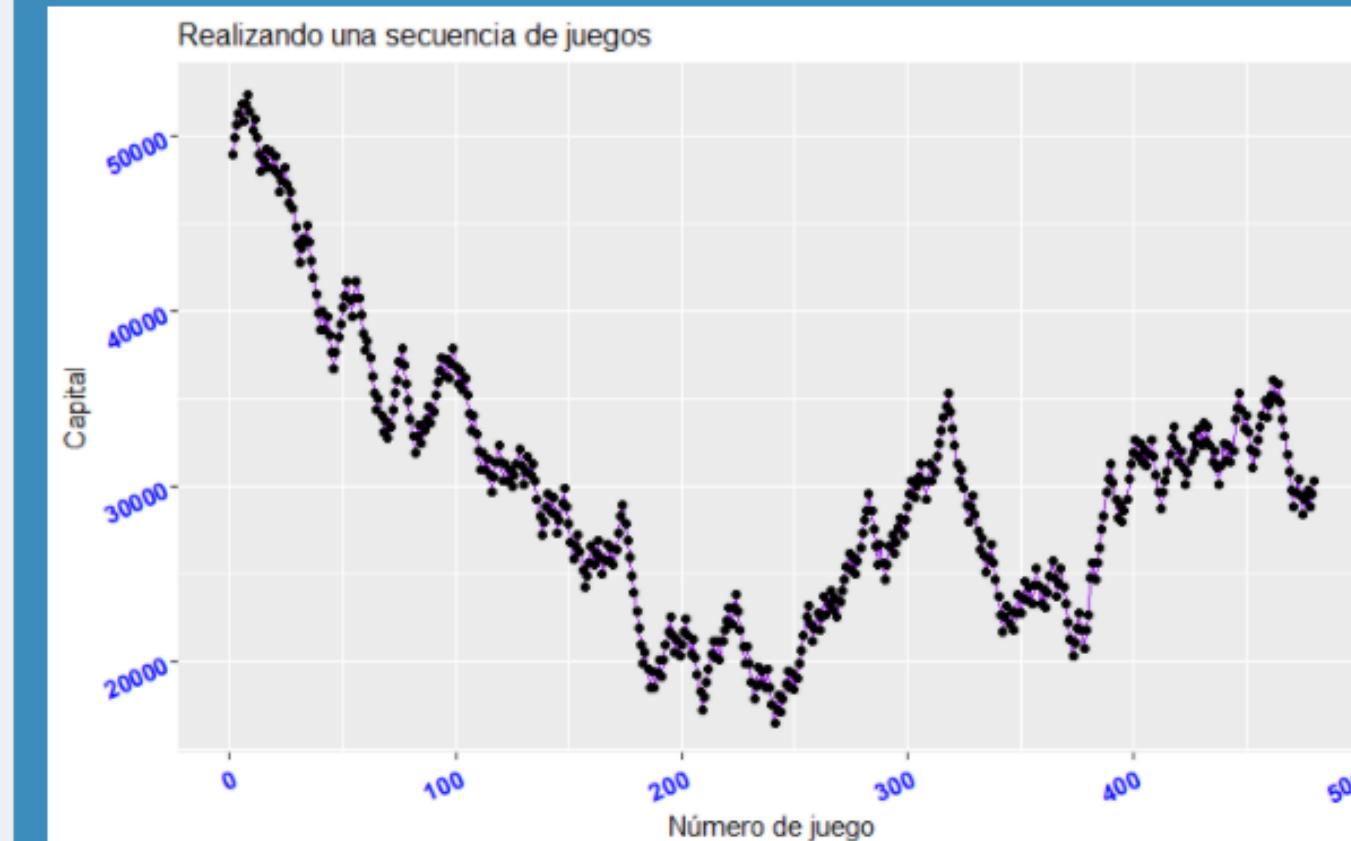
Postwork 08

Escenarios con momios máximos y promedio

Momios máximos



Momios promedio



Retos a
los que
enfrento
el equipo.



Repositorio

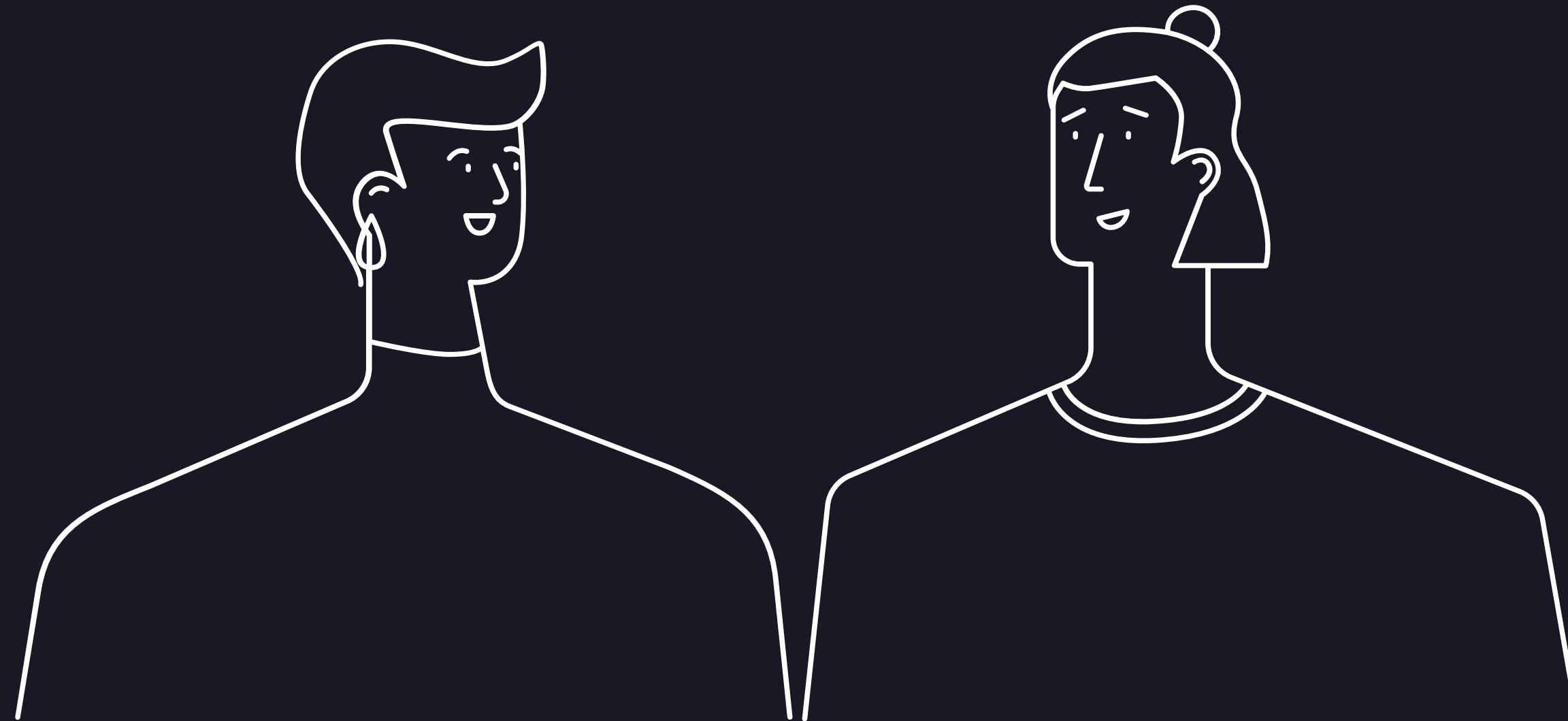


SCAN ME

https://github.com/gilesitorr/DataScience3_Bloque2

GRACIAS POR SU ATENCIÓN

B*EDU



Equipo 2