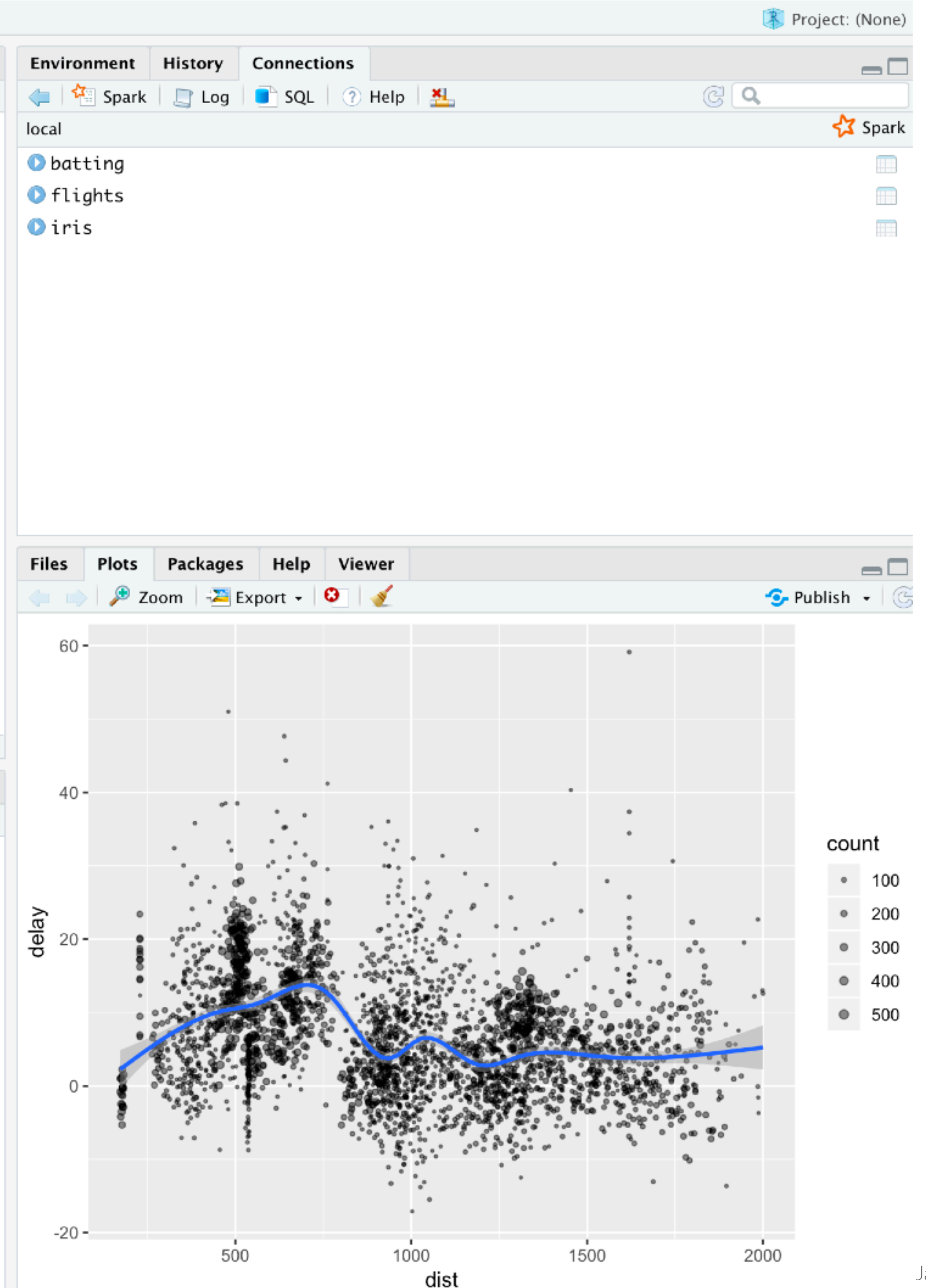


```
1 iris_tbl <- copy_to(sc, iris)
2 flights_tbl <- copy_to(sc, nycflights13::flights, "flights")
3 batting_tbl <- copy_to(sc, Lahman::Batting, "batting")
4 src_tbls(sc)
5
6 # filter by departure delay and print the first few records
7 flights_tbl %>% filter(dep_delay == 2)
8
9 delay <- flights_tbl %>%
10   group_by(tailnum) %>%
11   summarise(count = n(), dist = mean(distance), delay = mean(arr_delay)) %>%
12   filter(count > 20, dist < 2000, !is.na(delay)) %>%
13   collect
14
15 # plot delays
16 library(ggplot2)
17 ggplot(delay, aes(dist, delay)) +
18   geom_point(aes(size = count), alpha = 1/2) +
19   geom_smooth() +
20   scale_size_area(max_size = 2)
```

20:32 (Top Level) R Script

Console

```
> delay <- flights_tbl %>%
+   group_by(tailnum) %>%
+   summarise(count = n(), dist = mean(distance), delay = mean(arr_delay)) %>%
+   filter(count > 20, dist < 2000, !is.na(delay)) %>%
+   collect
Warnmeldung:
Missing values are always removed in SQL.
Use `mean(x, na.rm = TRUE)` to silence this warning
This warning is displayed only once per session.
>
> # plot delays
> library(ggplot2)
> ggplot(delay, aes(dist, delay)) +
+   geom_point(aes(size = count), alpha = 1/2) +
+   geom_smooth() +
+   scale_size_area(max_size = 2)
'geom_smooth()' using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



# Spark Jobs (?)

User: jankirenz

Total Uptime: 3,8 min

Scheduling Mode: FIFO

Active Jobs: 1

Completed Jobs: 7

▶ Event Timeline

## ▼ Active Jobs (1)

Job Id ▾	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
7	sql at NativeMethodAccessorImpl.java:0 sql at NativeMethodAccessorImpl.java:0 (kill)	2019/05/21 20:26:20	1 s	0/2	0/2

## ▼ Completed Jobs (7)

Job Id ▾	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
6	collect at utils.scala:44 collect at utils.scala:44	2019/05/21 20:26:19	37 ms	1/1	2/2
5	collect at utils.scala:204 collect at utils.scala:204	2019/05/21 20:26:19	0,3 s	2/2	2/2
4	sql at NativeMethodAccessorImpl.java:0 sql at NativeMethodAccessorImpl.java:0	2019/05/21 20:26:13	5 s	2/2	2/2
3	collect at utils.scala:44 collect at utils.scala:44	2019/05/21 20:26:07	17 ms	1/1	1/1
2	collect at utils.scala:204 collect at utils.scala:204	2019/05/21 20:26:06	39 ms	2/2	2/2
1	sql at NativeMethodAccessorImpl.java:0 sql at NativeMethodAccessorImpl.java:0	2019/05/21 20:26:05	0,4 s	2/2	2/2
0	collect at utils.scala:44 collect at utils.scala:44	2019/05/21 20:26:04	0,9 s	1/1	1/1