UNIVERSITY OF CALIFORNIA, BERKELEY

STAT 133 FALL 2024

Lecture Notes Taken in the Class

November 4th, 2024 – Monday

COURSE: MATH 133 – CONCEPTS IN COMPUTING WITH DATA

INSTRUCTOR: GASTON SANCHEZ

NAME: **ALDAN OU**

DATE: NOVEMBER 4TH, 2024

❖ GEOSPARTIAL

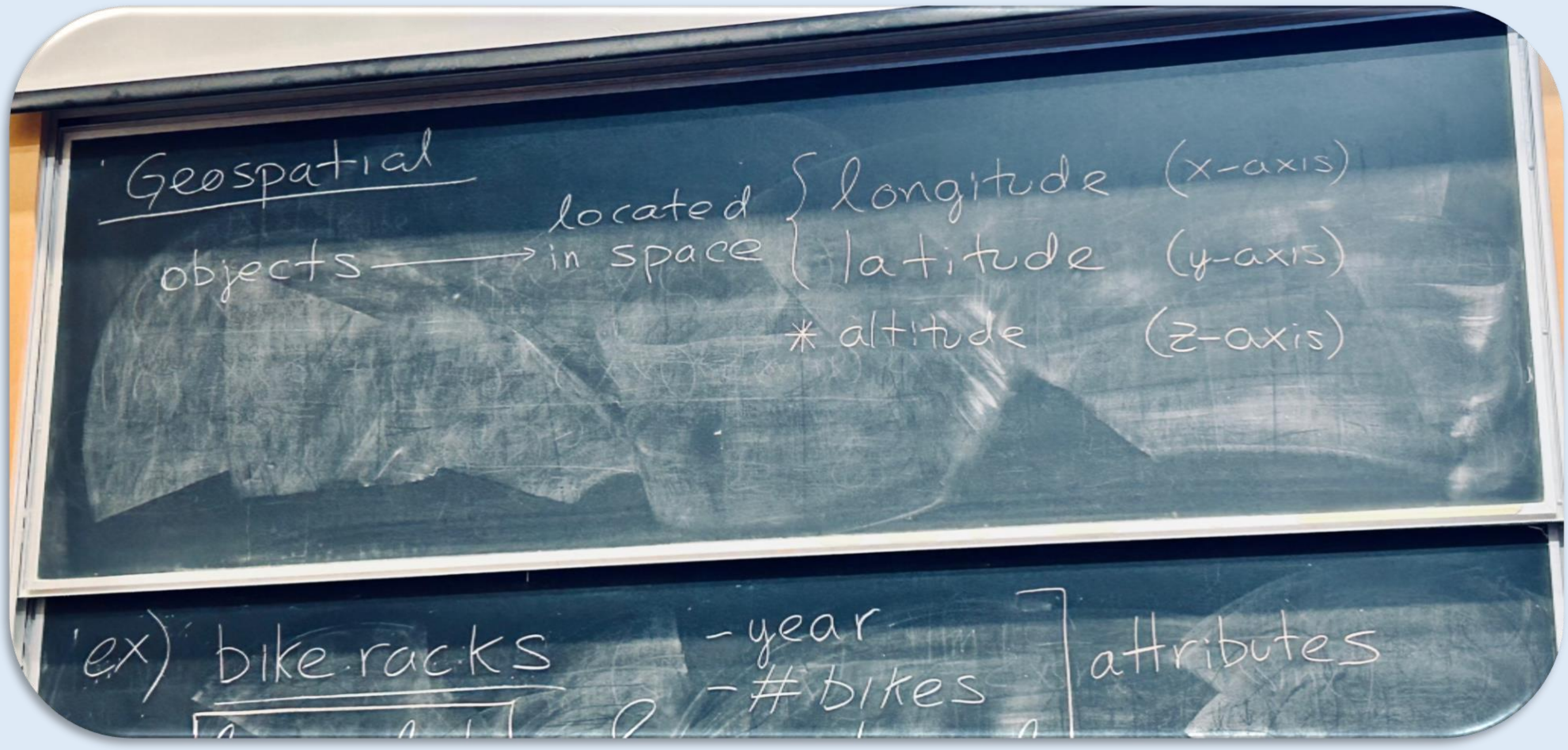
Geospartial refers to objects that are located in space using three primary coordinates.

- Longitude (x-axis): measures east-west position
- Latitude (y-axis): Measures north-south position
- Altitude (z-axis): Measures height / elevation

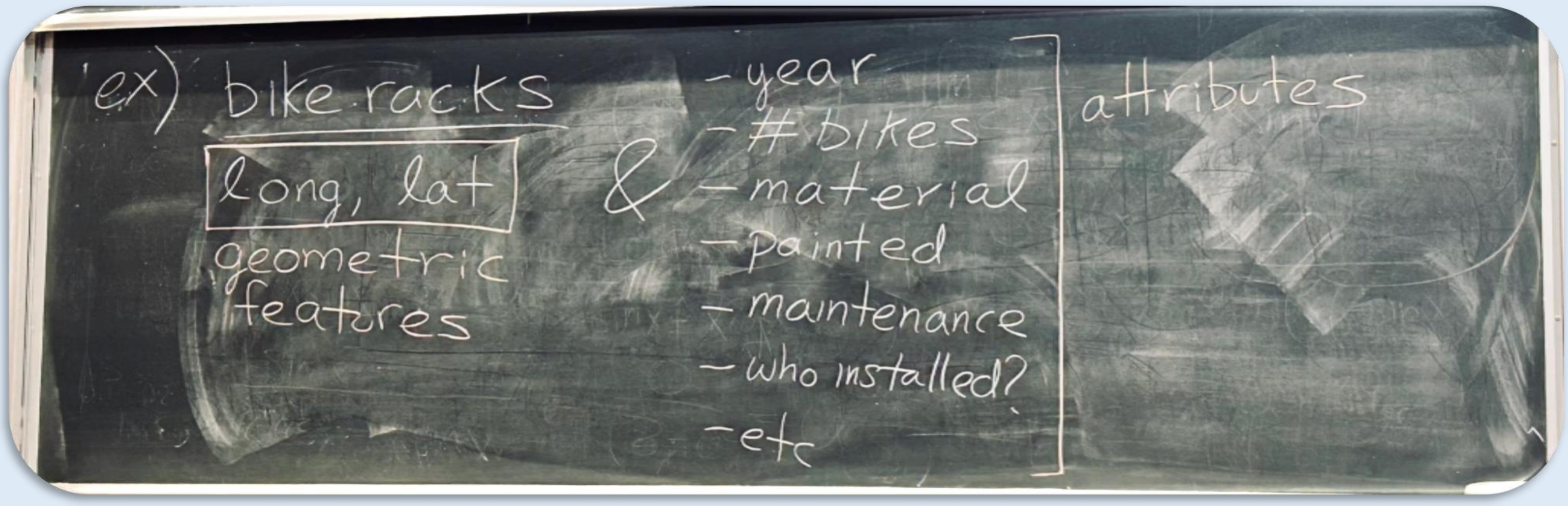
A common application of Geospartial is mapping storm and hurricane data.

1. Locations can be plotted using longitude and latitude coordinates;
2. Paths can be traced using connecting lines;
3. Data can be layered on different types of maps;
    1) Coastline maps
    2) Country maps
    3) Continental maps

The following photos are Prof. Sanchez's lecture notes that were taken in the class.



Example: Bike Racks

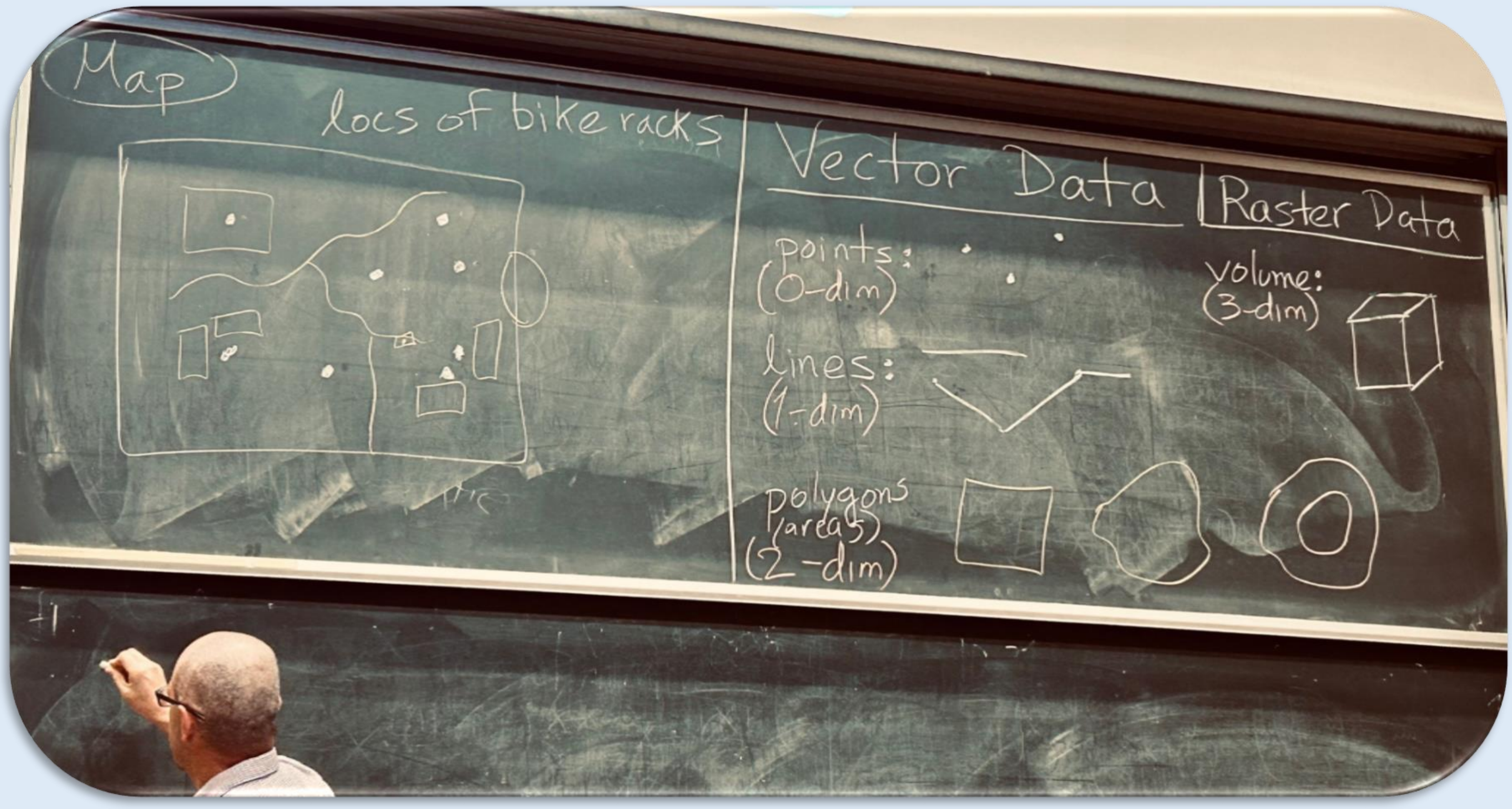


- How does this example illustrate the key aspects of geospatial data?

1) The first one is the **Spatial Component.** The geometric features (longitude and latitude) that specify where objects are located in physical space.
2) The second one is the **Descriptive Component.** The attributes that provide additional information about the object at those locations. Such attributes include several non-spatial characteristics: year of installation, number of bikes, material used, paint condition, and maintenance history…

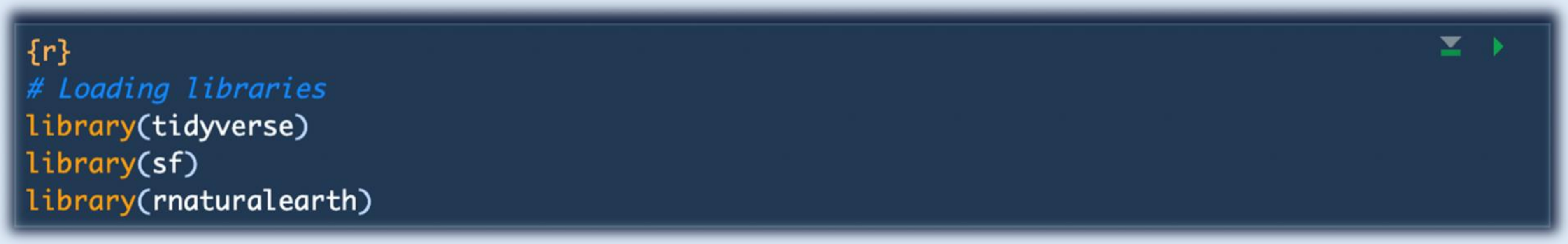
❖ VECTOR DATA

The following photo – Vector Data, is Prof. Sanchez's lecture notes that were taken in the class.



❖ R STUDIO CODING

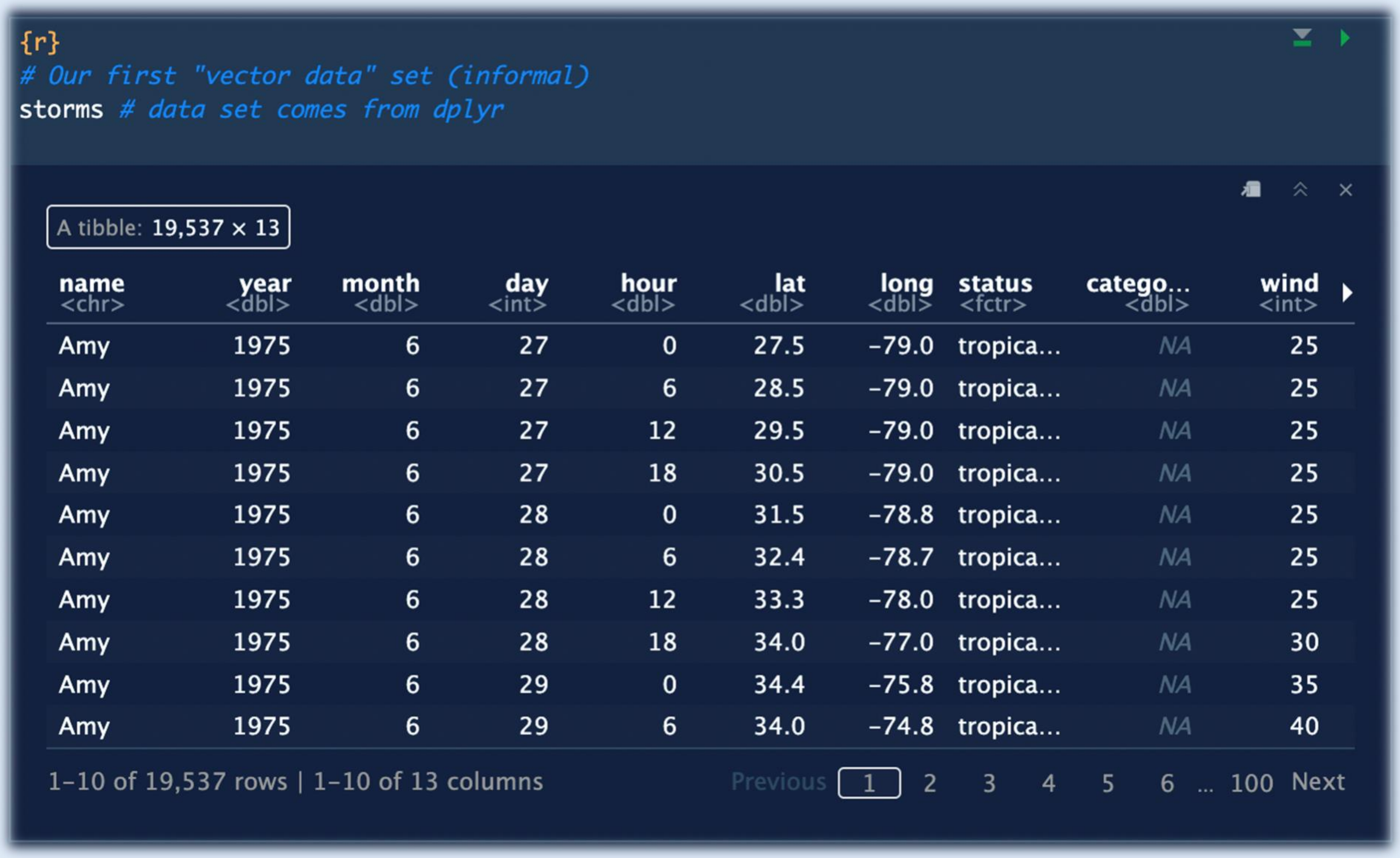
Now, it’s time to do some R coding.



```
{r}
# Loading libraries
library(tidyverse)
library(sf)
library(rnaturalearth)
```

The R package “sf” provides powerful tools for handling geospatial data with several advanges.

- Fast reading and writing of geospatial data
- Enhanced plotting performance
- Compatibility with data frames
- Consistent function naming

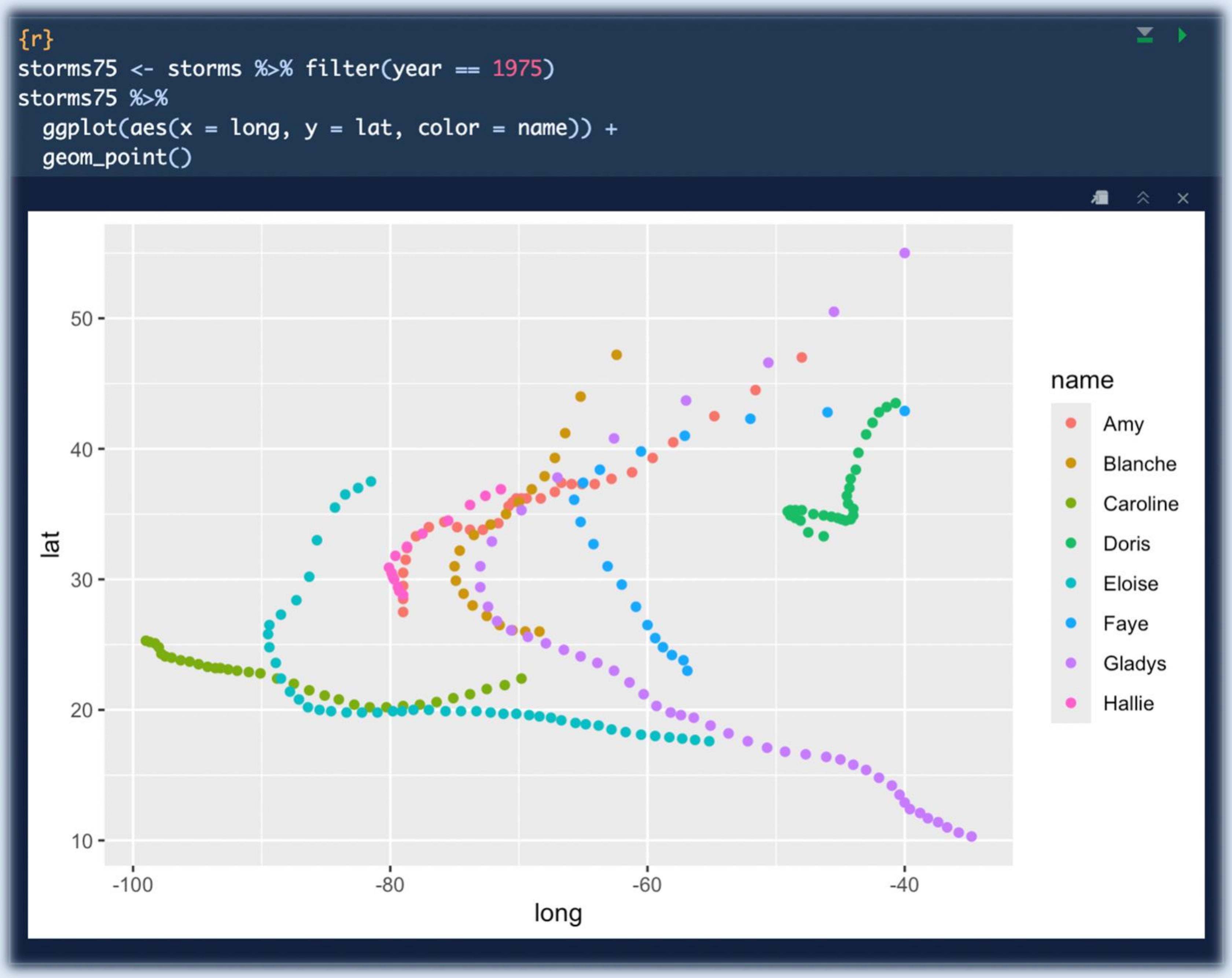


```
{r}
# Our first "vector data" set (informal)
storms # data set comes from dplyr
```

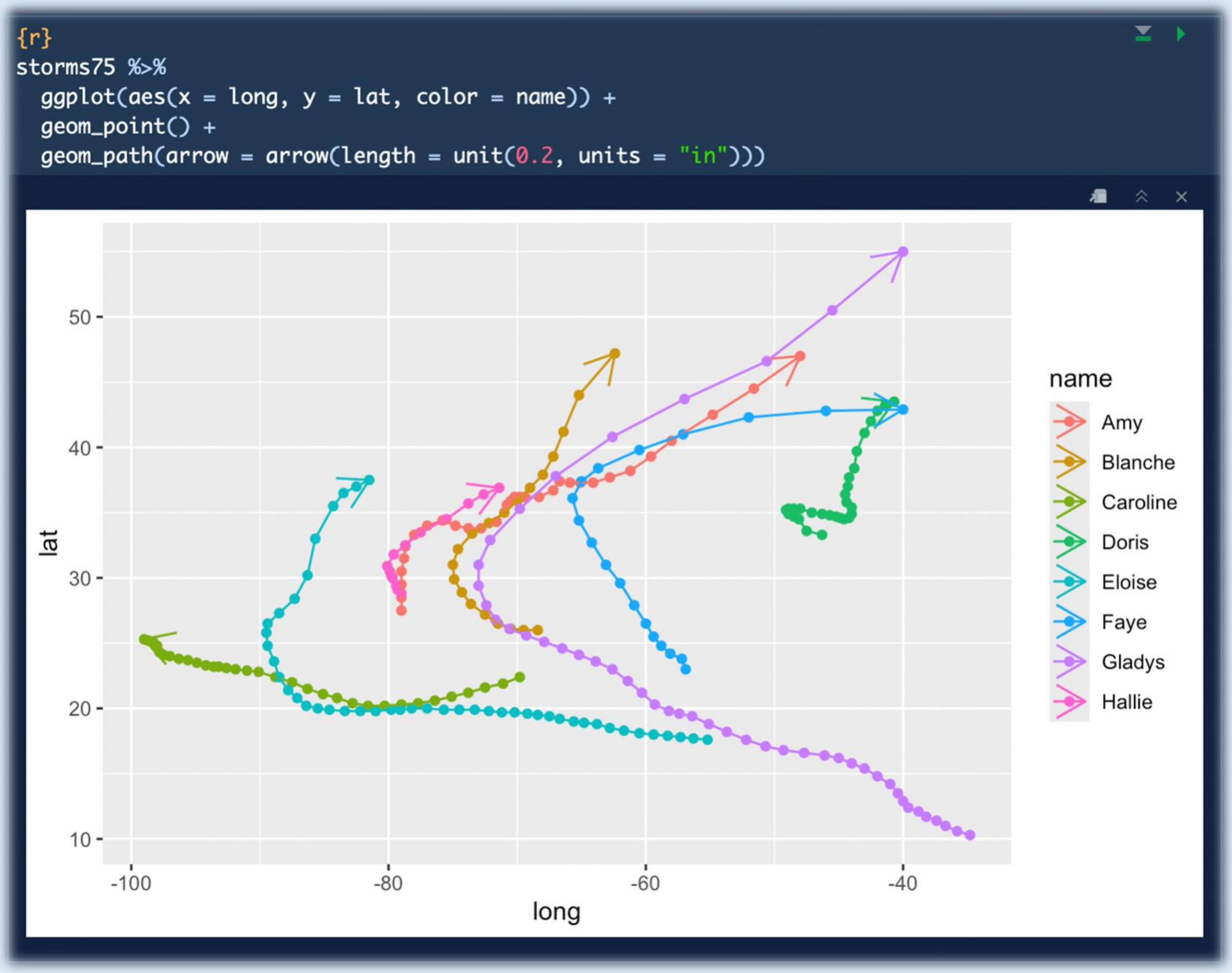
```
A tibble: 19,537 × 13

name    year   month   day   hour   lat   long   status     catego…  wind
<chr>   <dbl>  <dbl>   <int> <dbl>  <dbl> <dbl>  <fctr>     <dbl>    <int>
Amy     1975   6       27    0      27.5  -79.0  tropica…   NA       25
Amy     1975   6       27    6      28.5  -79.0  tropica…   NA       25
Amy     1975   6       27    12     29.5  -79.0  tropica…   NA       25
Amy     1975   6       27    18     30.5  -79.0  tropica…   NA       25
Amy     1975   6       28    0      31.5  -78.8  tropica…   NA       25
Amy     1975   6       28    6      32.4  -78.7  tropica…   NA       25
Amy     1975   6       28    12     33.3  -78.0  tropica…   NA       25
Amy     1975   6       28    18     34.0  -77.0  tropica…   NA       30
Amy     1975   6       29    0      34.4  -75.8  tropica…   NA       35
Amy     1975   6       29    6      34.0  -74.8  tropica…   NA       40

1–10 of 19,537 rows | 1–10 of 13 columns        Previous  1  2  3  4  5  6  ...  100  Next
```

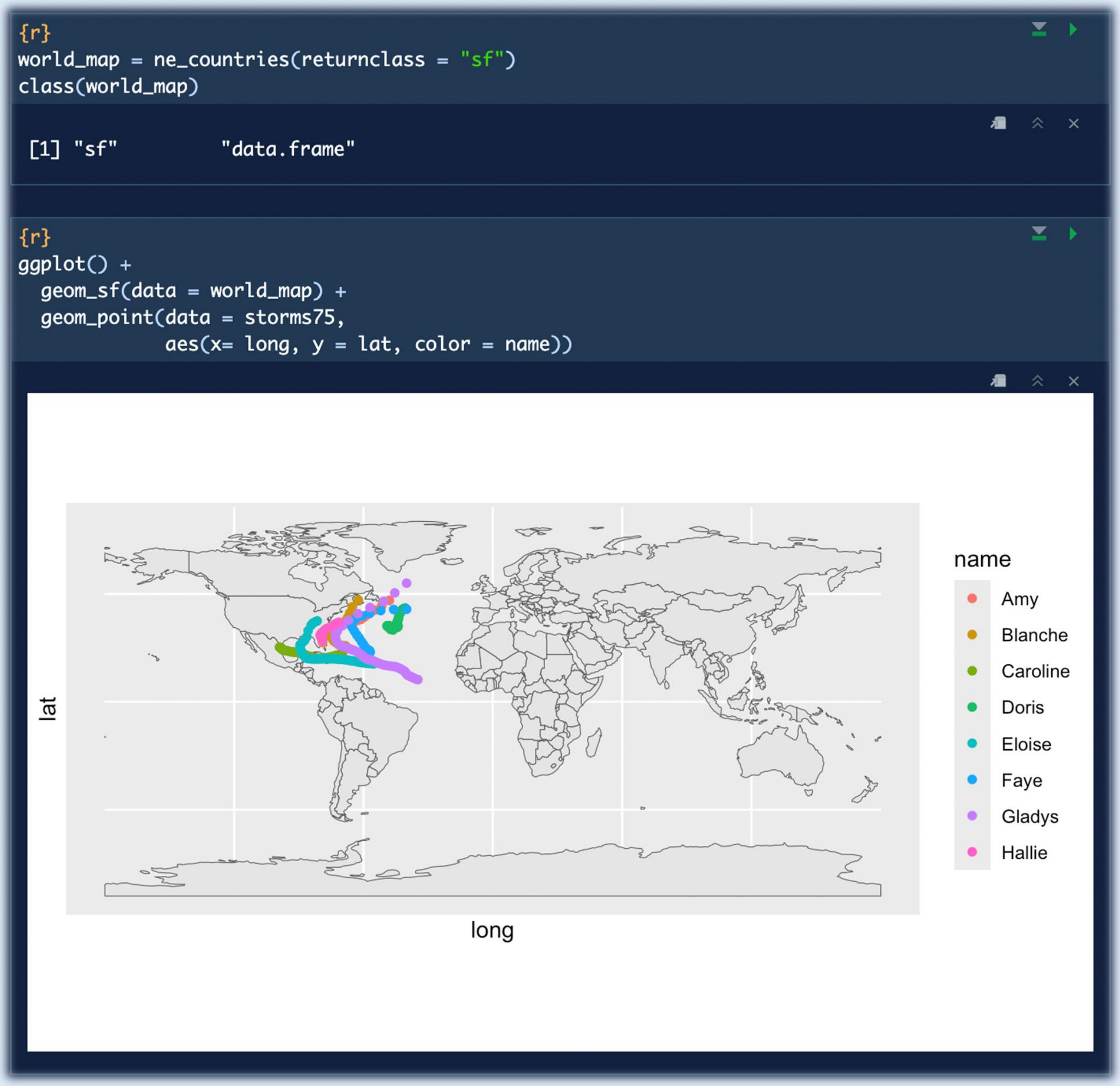


```
{r}
storms75 <- storms %>% filter(year == 1975)
storms75 %>%
  ggplot(aes(x = long, y = lat, color = name)) +
  geom_point()
```



```
{r}
storms75 %>%
  ggplot(aes(x = long, y = lat, color = name)) +
  geom_point() +
  geom_path(arrow = arrow(length = unit(0.2, units = "in")))
```

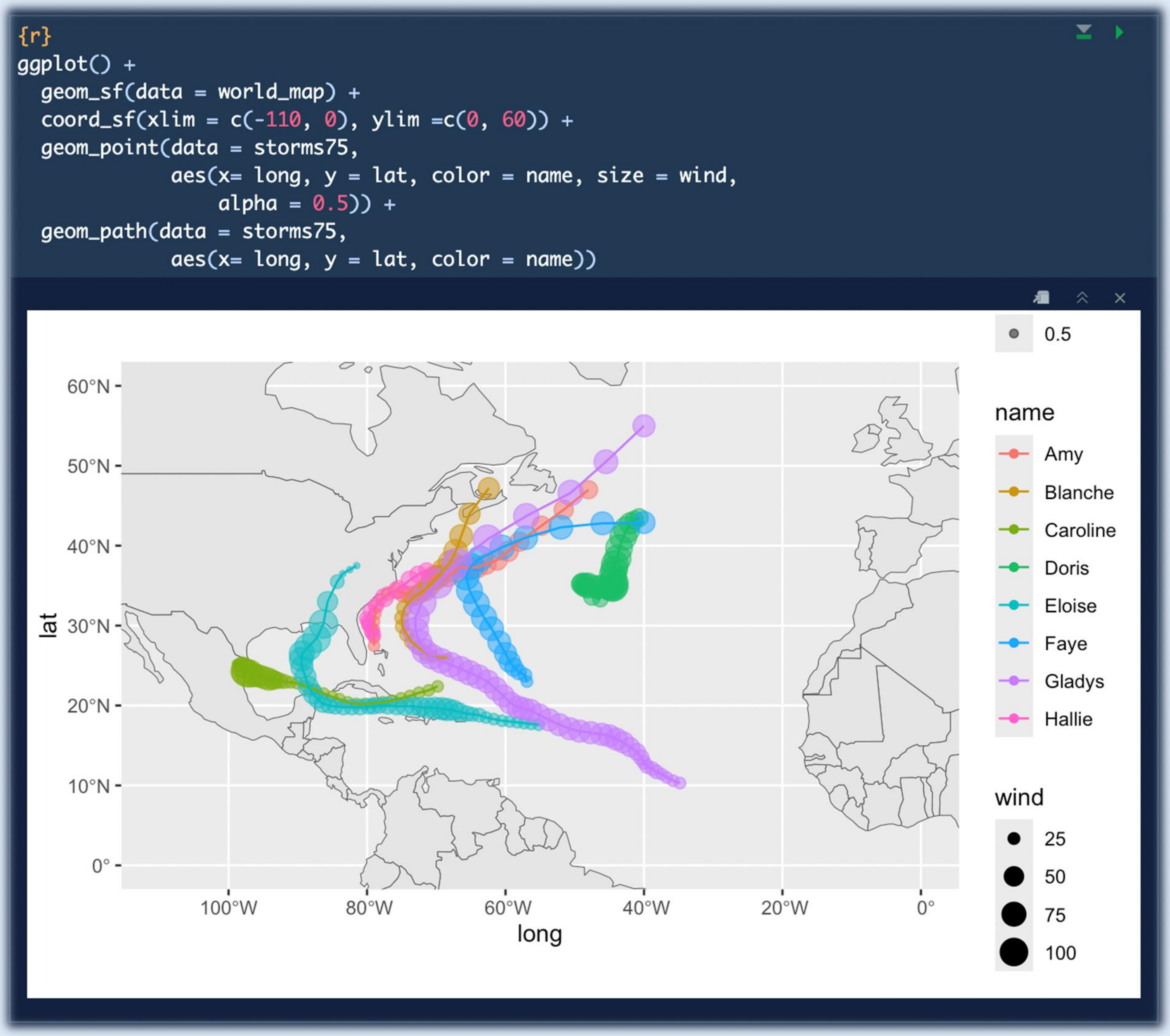
❖ MAPS



```
{r}
world_map = ne_countries(returnclass = "sf")
class(world_map)
```

```
[1] "sf"          "data.frame"
```

```
{r}
ggplot() +
  geom_sf(data = world_map) +
  geom_point(data = storms75,
             aes(x= long, y = lat, color = name))
```



```
{r}
ggplot() +
  geom_sf(data = world_map) +
  coord_sf(xlim = c(-110, 0), ylim =c(0, 60)) +
  geom_point(data = storms75,
             aes(x= long, y = lat, color = name, size = wind,
                 alpha = 0.5)) +
  geom_path(data = storms75,
             aes(x= long, y = lat, color = name))
```