Vis I

Import the weather data

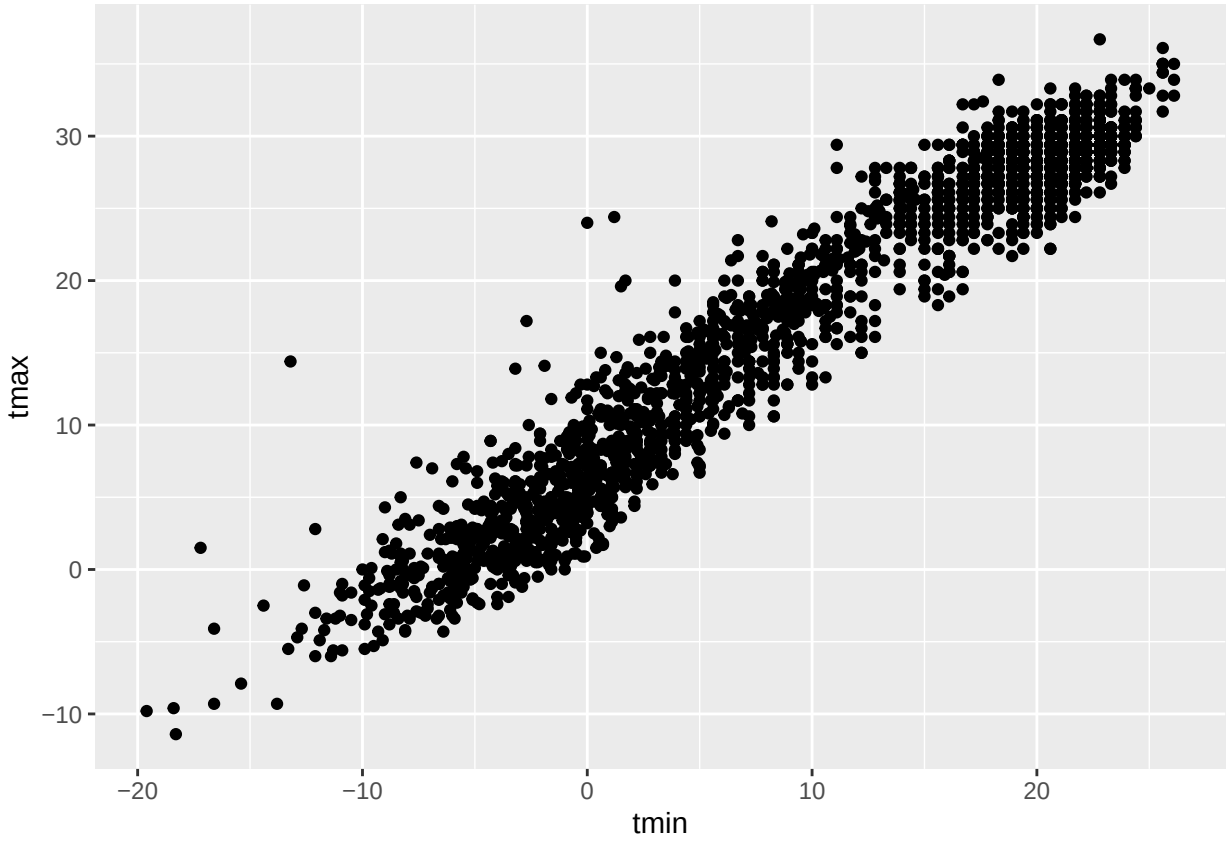
```
weather_df =
  rnoaa::meteo_pull_monitors(
    c("USW00094728", "USW00022534", "USS0023B17S"),
    var = c("PRCP", "TMIN", "TMAX"),
    date_min = "2021-01-01",
    date_max = "2022-12-31") |>
  mutate(
    name = case_match(
      id,
      "USW00094728" ~ "CentralPark_NY",
      "USW00022534" ~ "Molokai_HI",
      "USS0023B17S" ~ "Waterhole_WA"),
    tmin = tmin / 10,
    tmax = tmax / 10) |>
  select(name, id, everything())
```

```
## using cached file: /Users/soomin.you/Library/Caches/org.R-project.R/R/rnoaa/noaa_ghcnd/USW00094728.d

## date created (size, mb): 2024-09-03 14:09:15.067935 (8.636)

## file min/max dates: 1869-01-01 / 2024-09-30

## using cached file: /Users/soomin.you/Library/Caches/org.R-project.R/R/rnoaa/noaa_ghcnd/USW00022534.d

## date created (size, mb): 2024-09-03 14:09:24.583853 (3.913)

## file min/max dates: 1949-10-01 / 2024-09-30

## using cached file: /Users/soomin.you/Library/Caches/org.R-project.R/R/rnoaa/noaa_ghcnd/USS0023B17S.d

## date created (size, mb): 2024-09-03 14:09:27.654133 (1.036)

## file min/max dates: 1999-09-01 / 2024-08-31
```

Making out first plot

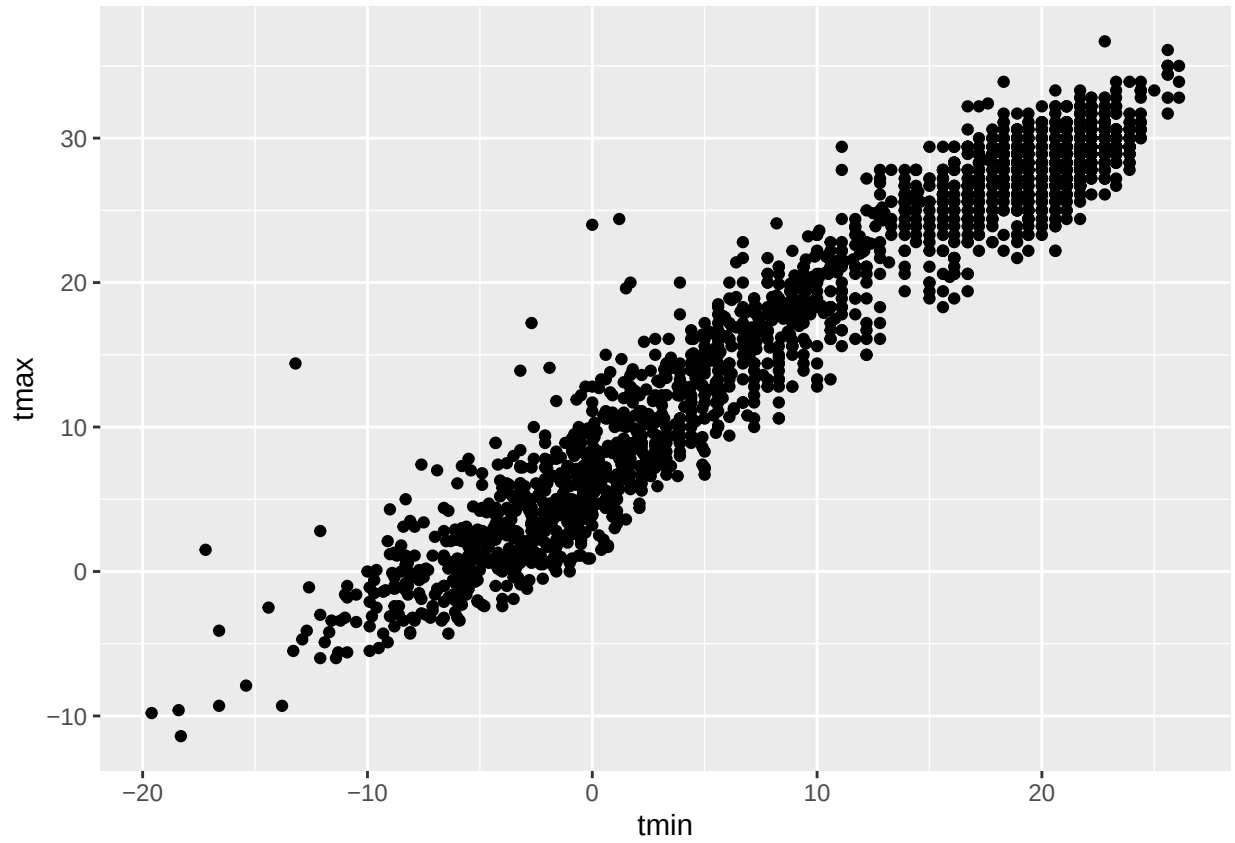
```
ggplot(weather_df, aes(x = tmin, y = tmax)) +
  geom_point()
```

```
## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



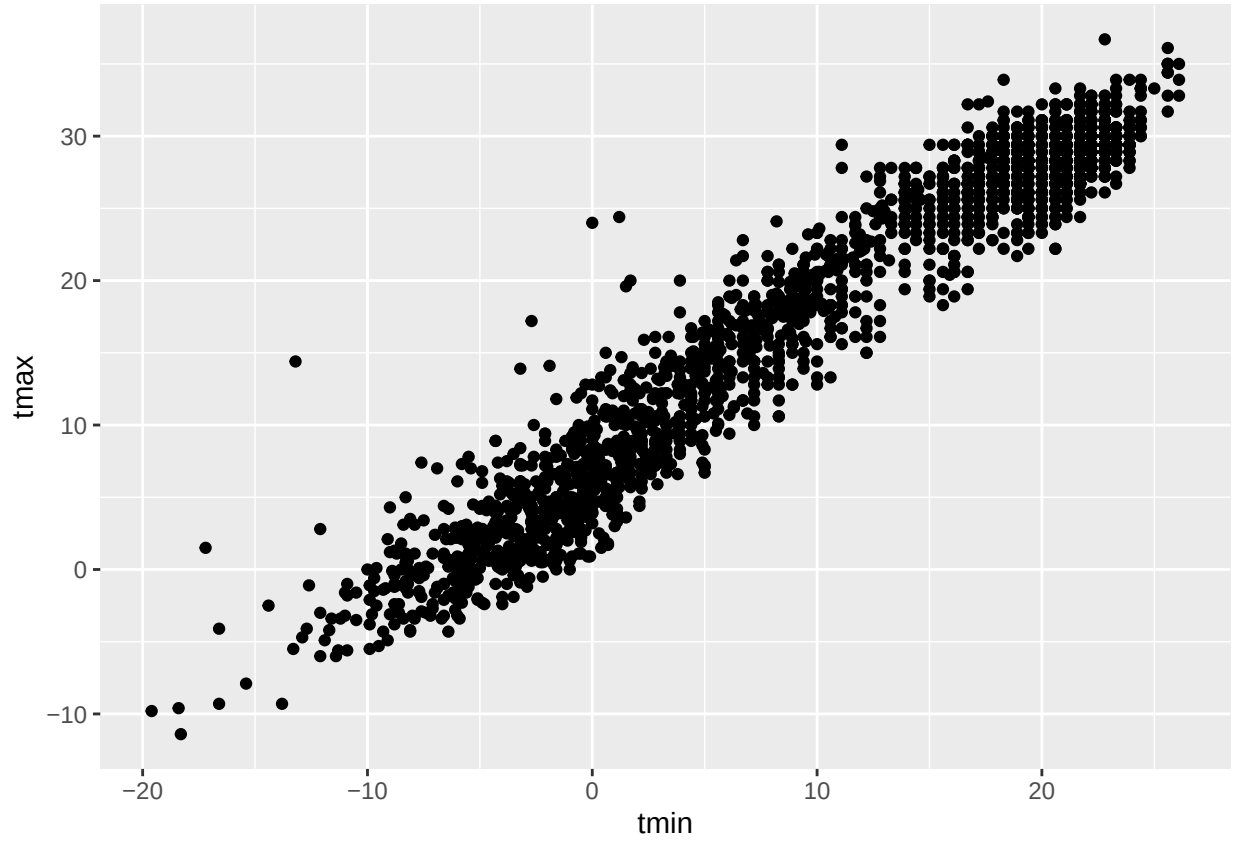

```
weather_df |>
  ggplot(aes(x = tmin, y = tmax)) +
  geom_point()
```

```
## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```




```
ggp_weather_scatterplot =
  weather_df |>
  ggplot(aes(x = tmin, y = tmax)) +
  geom_point()

ggp_weather_scatterplot
```

```
## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



Check why some are missing

CHECK THE RECORDING

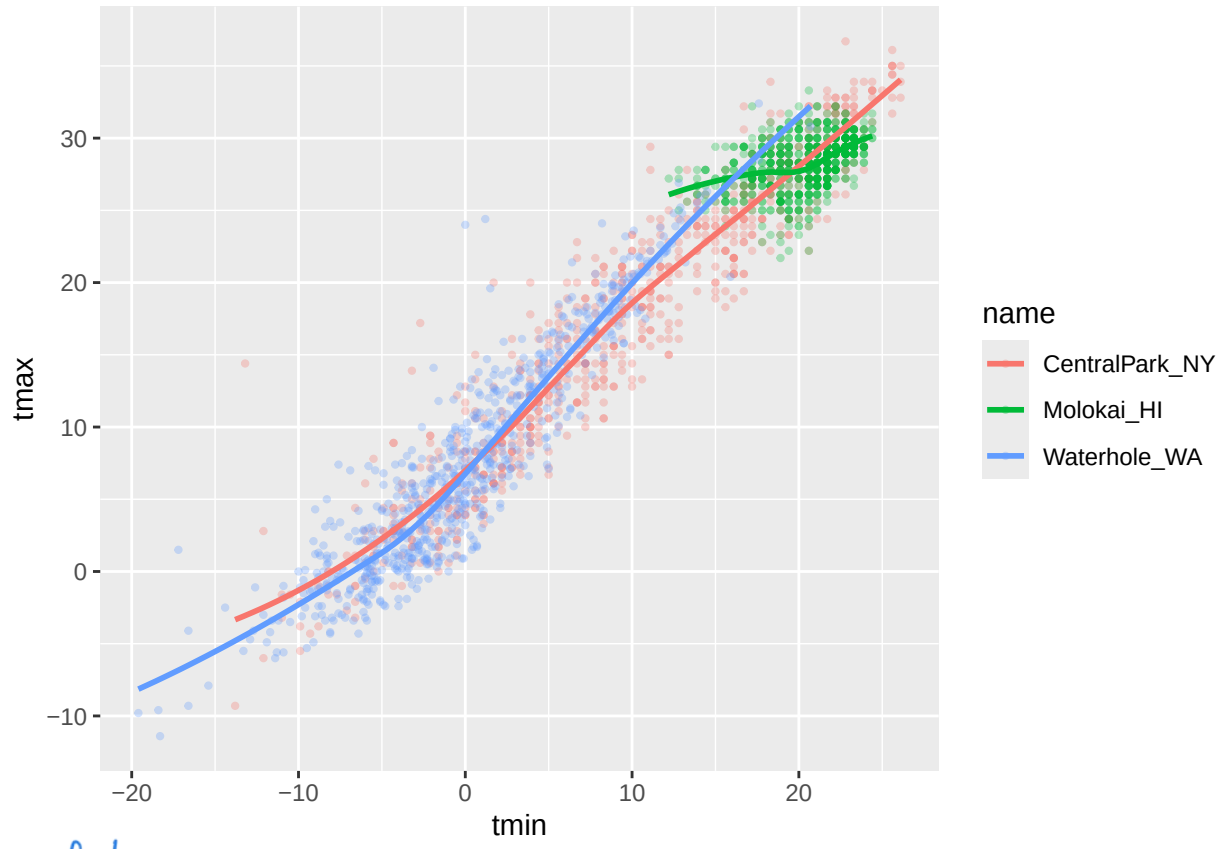
##Fancier scatterplots!

```
weather_df |>
  ggplot(aes(x = tmin, y = tmax, color = name)) +
  geom_point(alpha = 0.3, size = 0.8) +
  geom_smooth(se = FALSE)
```

size

transparency

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'

## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



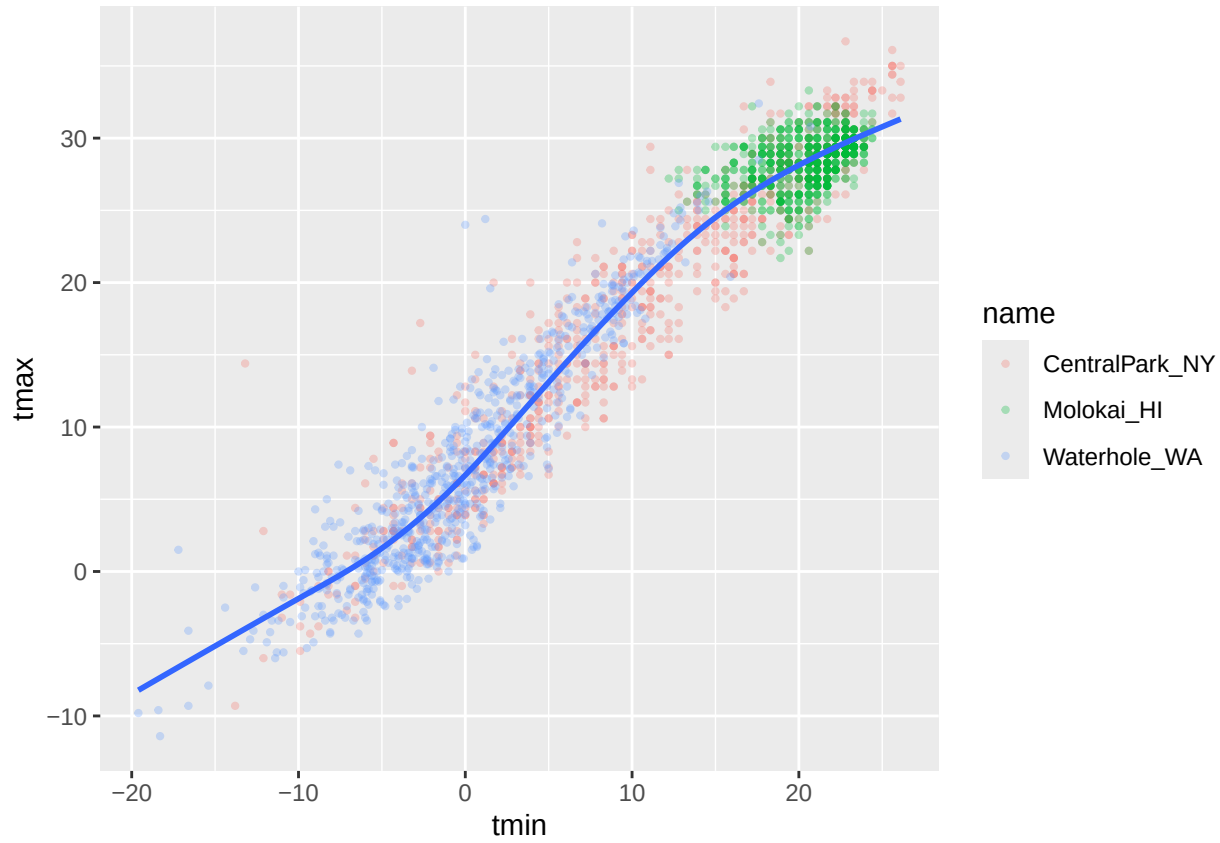
Where you define aesthetics can matter

```
weather_df |>
  ggplot(aes(x = tmin, y = tmax)) +
  geom_point(aes(color = name), alpha = 0.3, size = 0.8) +
  geom_smooth(se = FALSE)
```

```
## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```

```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_smooth()`).
```

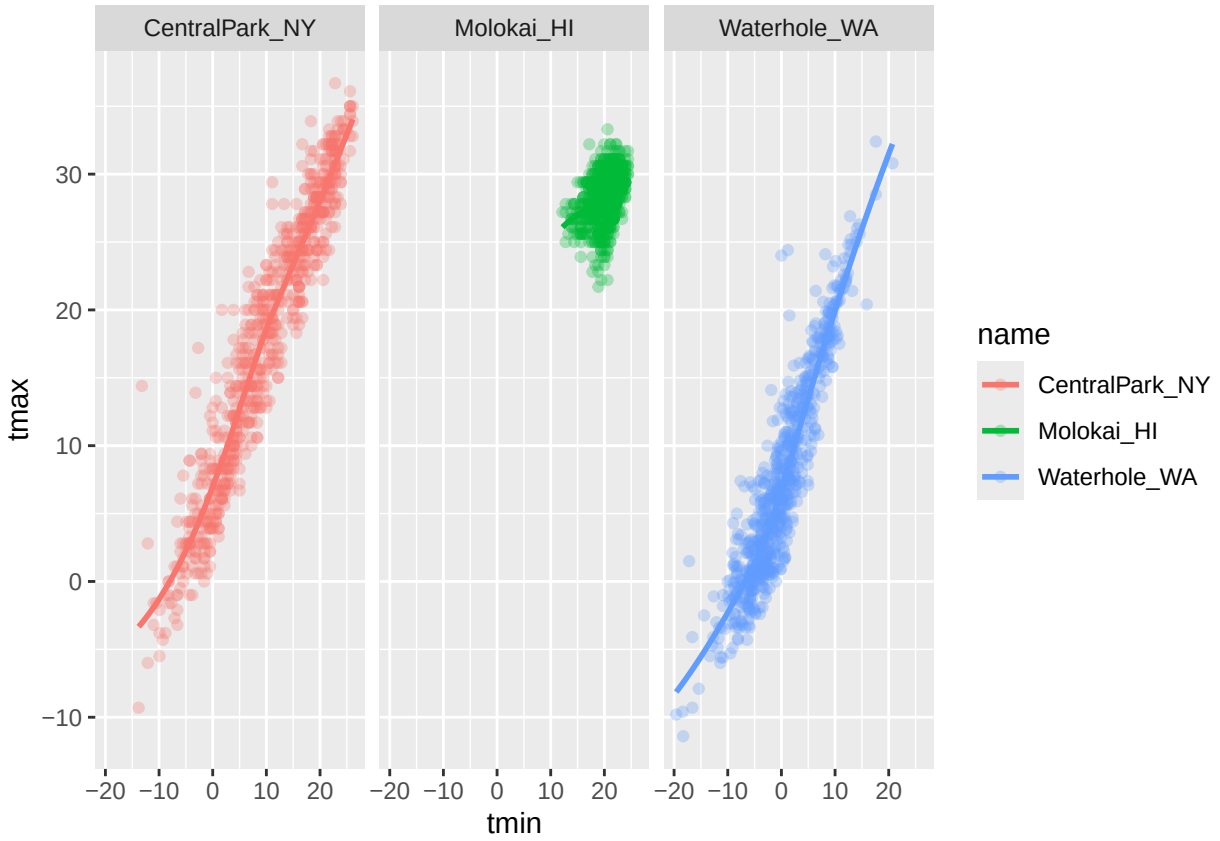
```
## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



use faceting real quick

```
weather_df |>
  ggplot(aes(x = tmin, y = tmax, color = name)) +
  geom_point(alpha = 0.3) +
  geom_smooth(se = FALSE) +
  facet_grid(. ~ name)
```

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'

## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_smooth()`).

## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

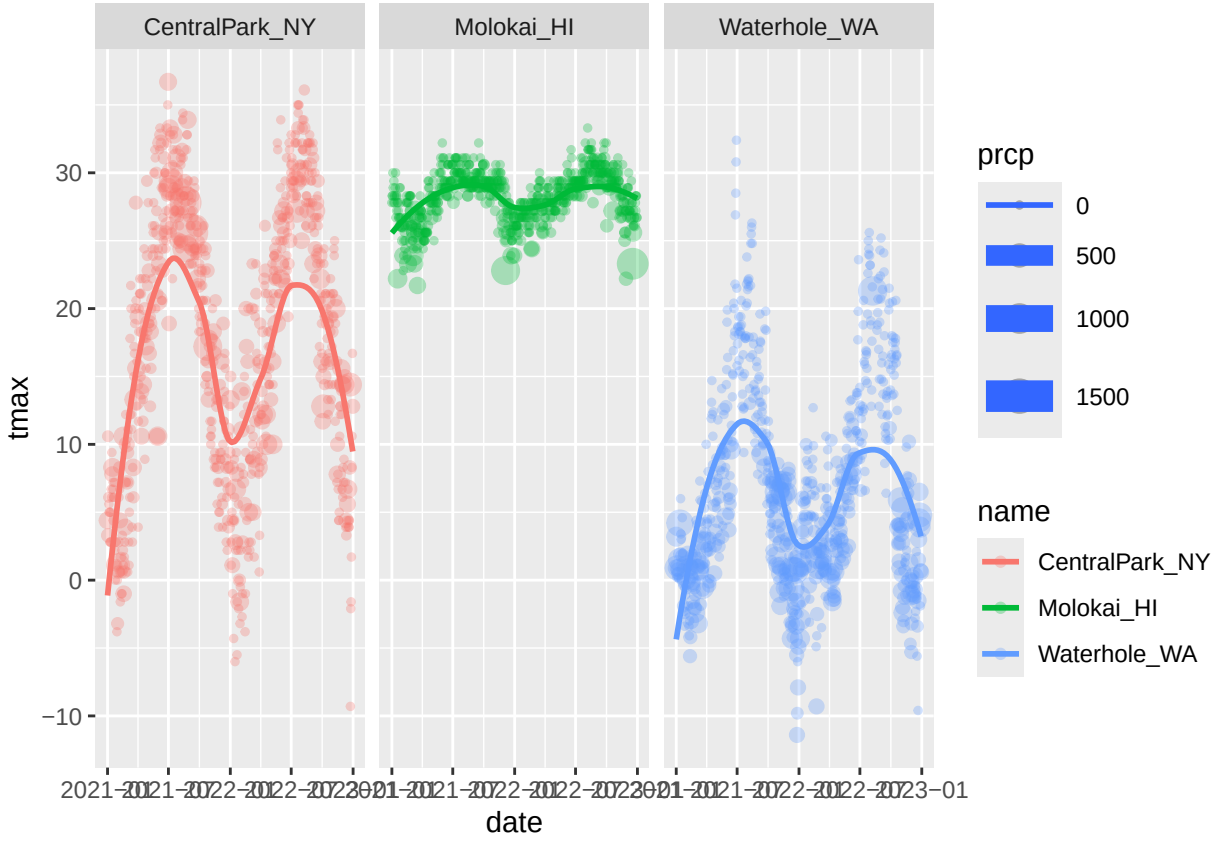


Let's make a somewhat more interesting scatterplot

```
weather_df |>
  ggplot(aes(x = date, y = tmax, color = name, size = prcp)) +
  geom_point(alpha = 0.3) +
  geom_smooth(se = FALSE) +
  facet_grid(. ~ name)
```

```
## Warning: Using ‘size‘ aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use ‘linewidth‘ instead.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.

## ‘geom_smooth()‘ using method = ’loess’ and formula = ’y ~ x’

## Warning: Removed 17 rows containing non-finite outside the scale range
## (‘stat_smooth()‘).

## Warning: The following aesthetics were dropped during statistical transformation: size.
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a ‘group‘ aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: size.
```

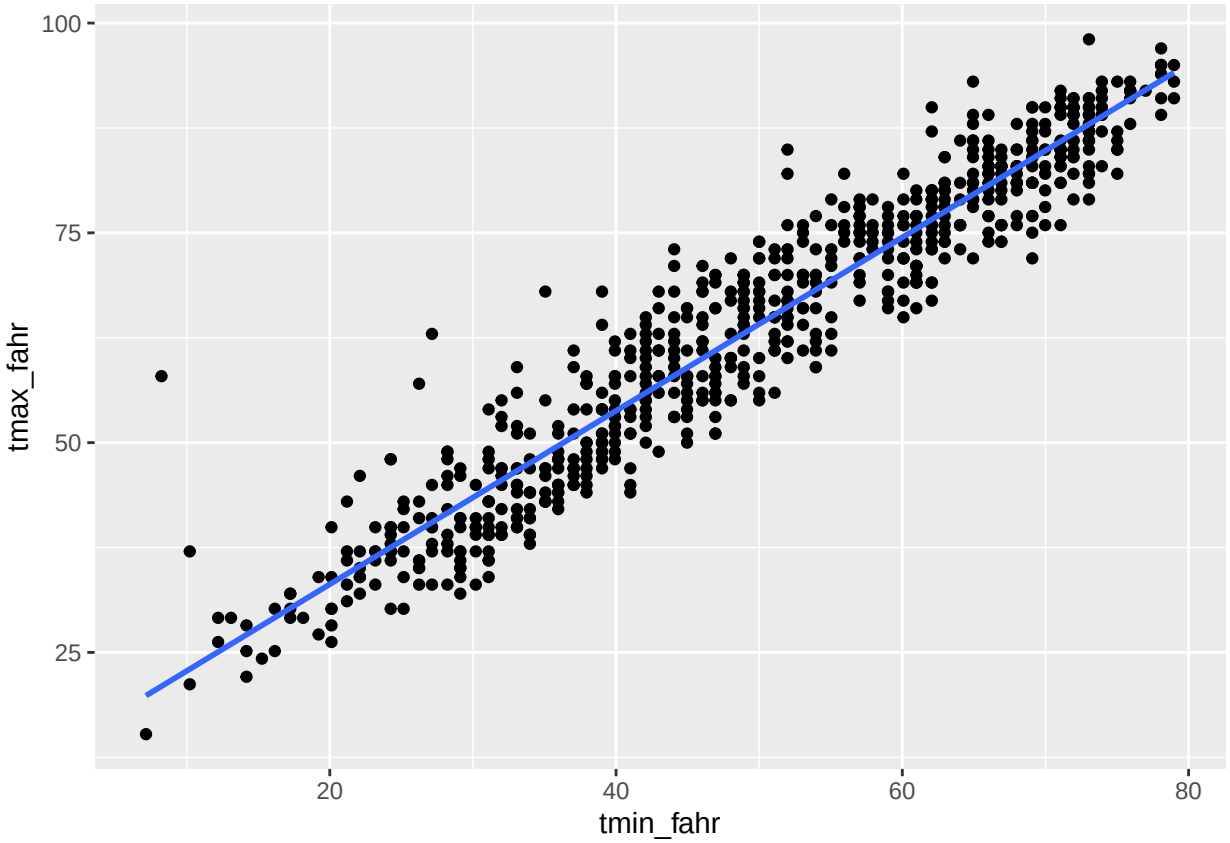
```
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?
## The following aesthetics were dropped during statistical transformation: size.
## i This can happen when ggplot fails to infer the correct grouping structure in
##   the data.
## i Did you forget to specify a `group` aesthetic or to convert a numerical
##   variable into a factor?

## Warning: Removed 19 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



Learning Assessment plot

```
weather_df |>
  filter(name == "CentralPark_NY") |>
  mutate(
    tmax_fahr = tmax * (9/5) + 32,
    tmin_fahr = tmin * (9/5) + 32
  ) |>
  ggplot(aes(x = tmin_fahr, y = tmax_fahr)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



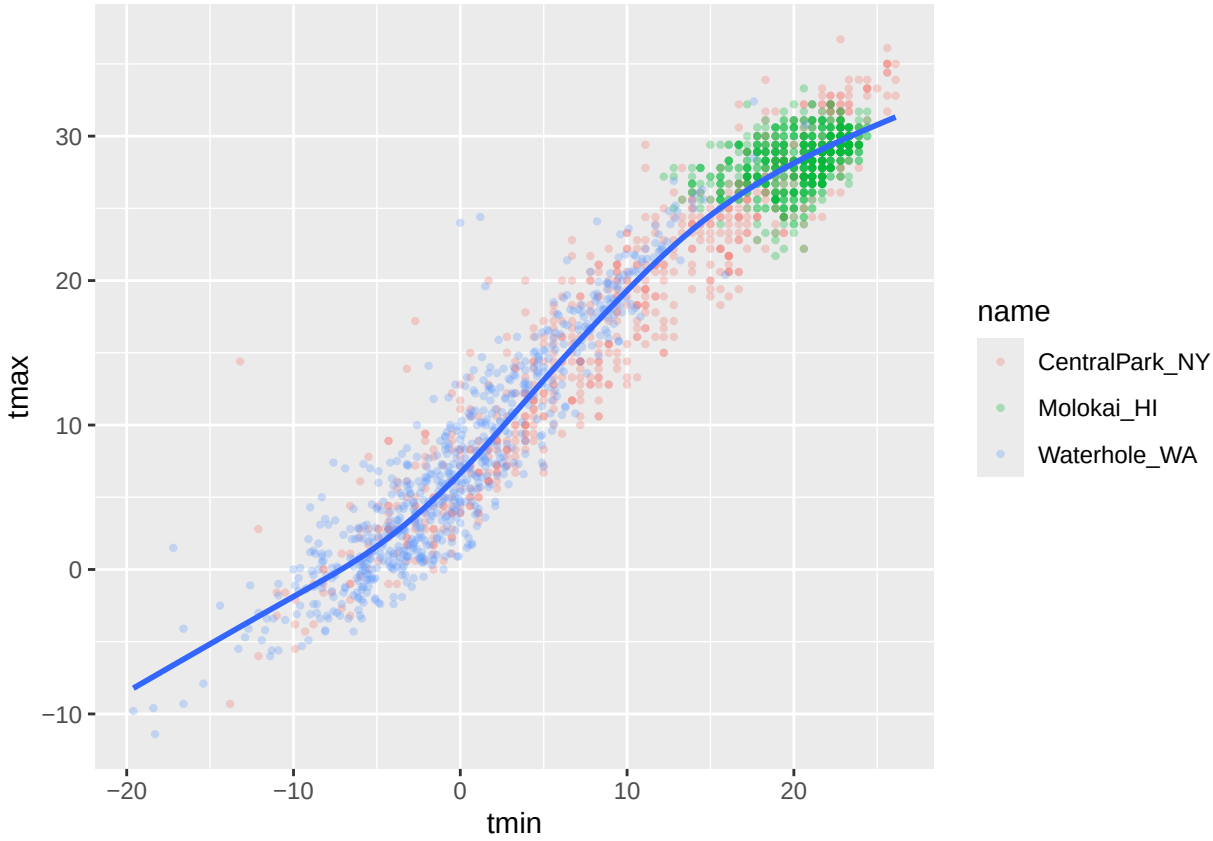
Small things

```
weather_df |>
  ggplot(aes(x = tmin, y = tmax)) +
  geom_point(aes(color = name), alpha = .3, size = .8) +
  geom_smooth(se = FALSE)
```

```
## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```

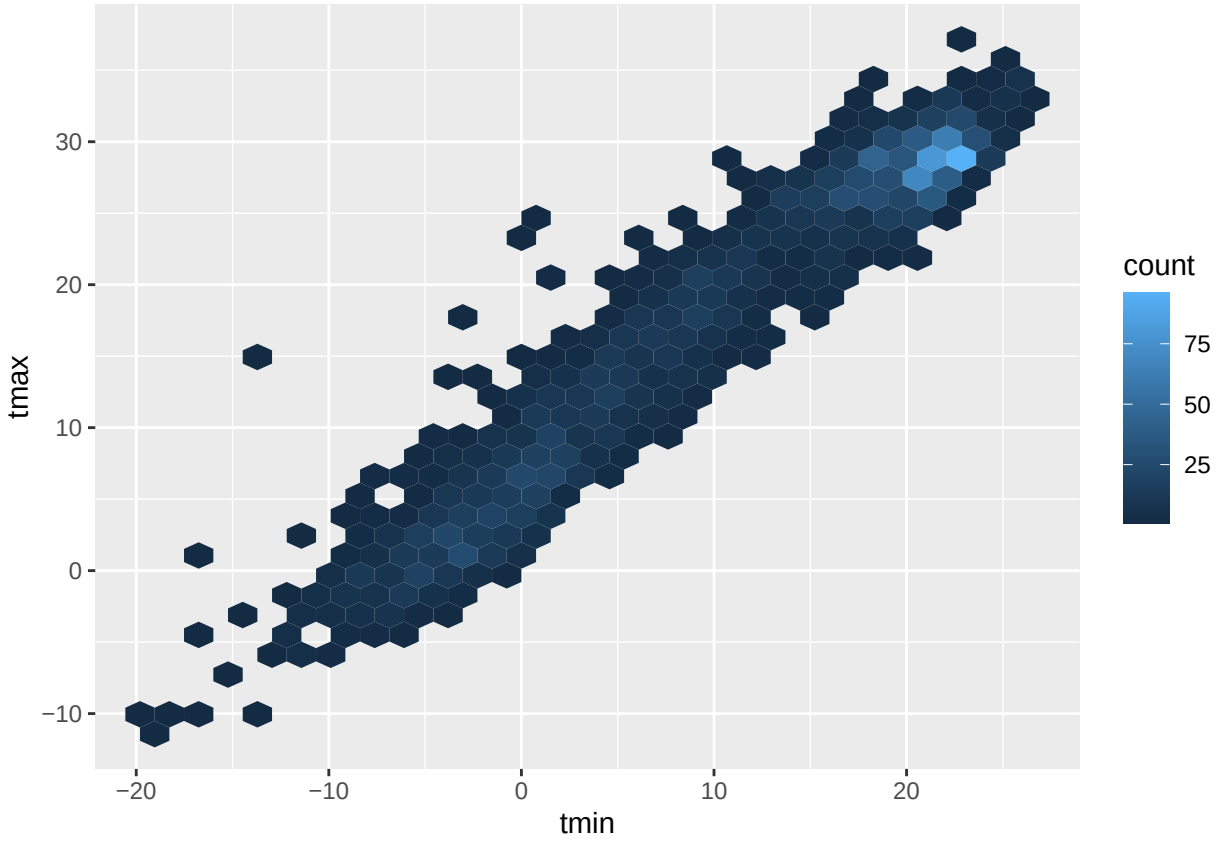
```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_smooth()`).
```

```
## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```



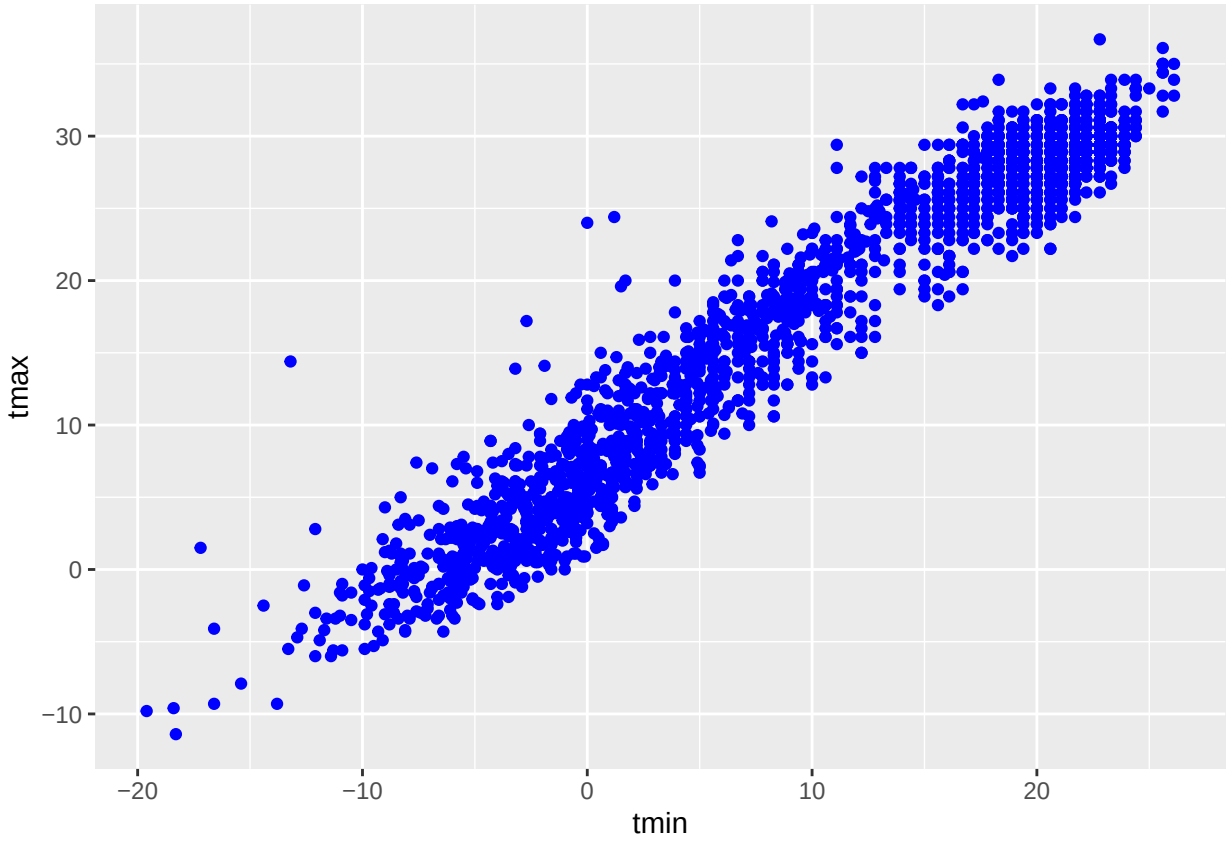

```
weather_df |>
  ggplot(aes(x = tmin, y = tmax)) +
  geom_hex()
```

```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_binhex()`).
```




```
weather_df |>
  ggplot(aes(x = tmin, y = tmax)) +
  geom_point(color = "blue")
```

```
## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

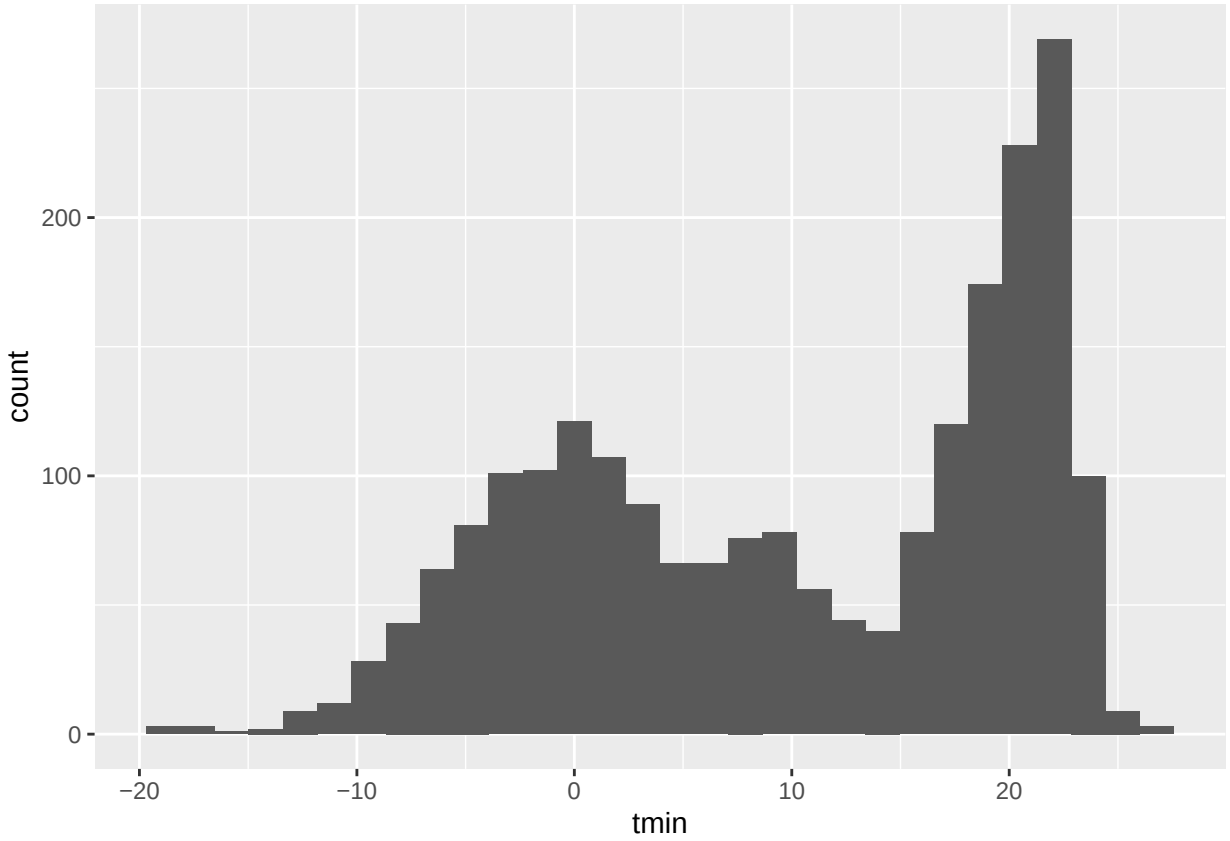


Univariate plots

```
weather_df |>
  ggplot(aes( x = tmin)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

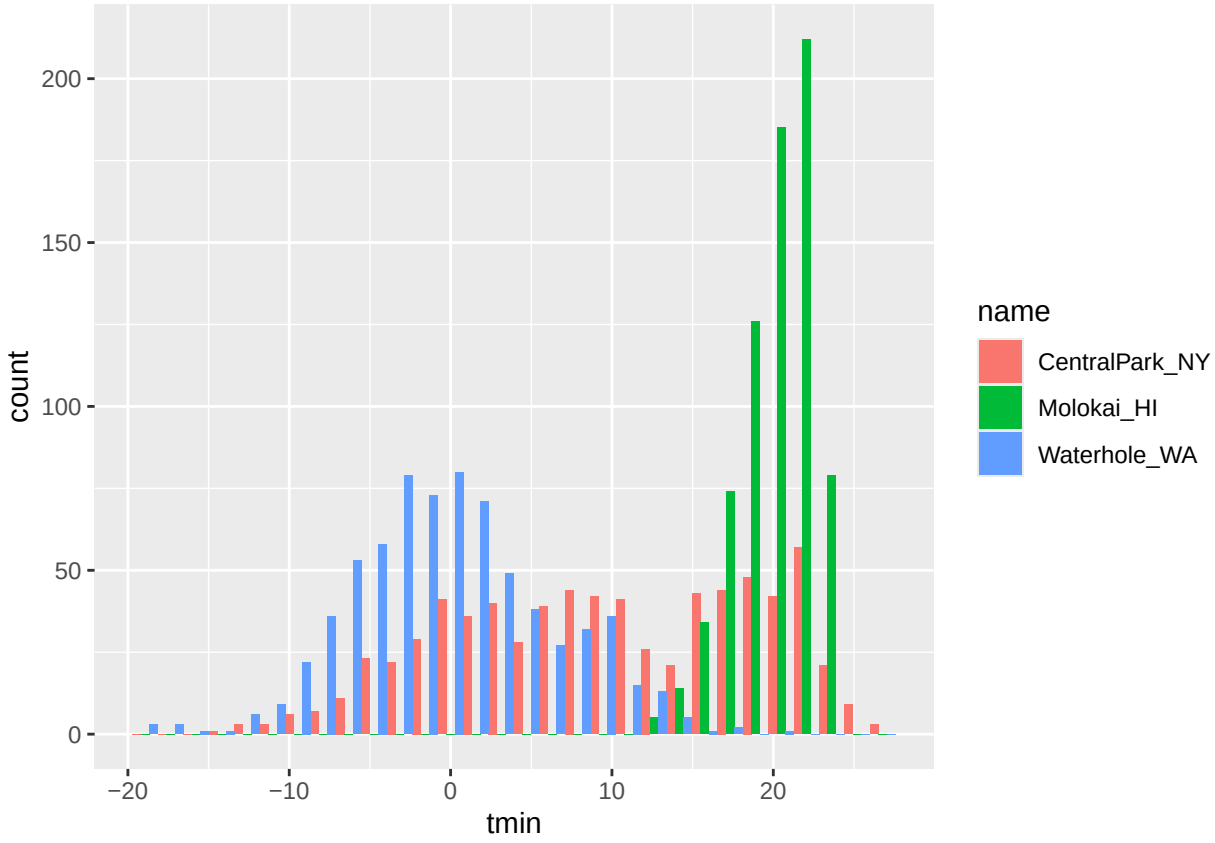
```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_bin()`).
```




```
weather_df |>
  ggplot(aes( x = tmin, fill = name)) +
  geom_histogram(position = "dodge")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_bin()`).
```

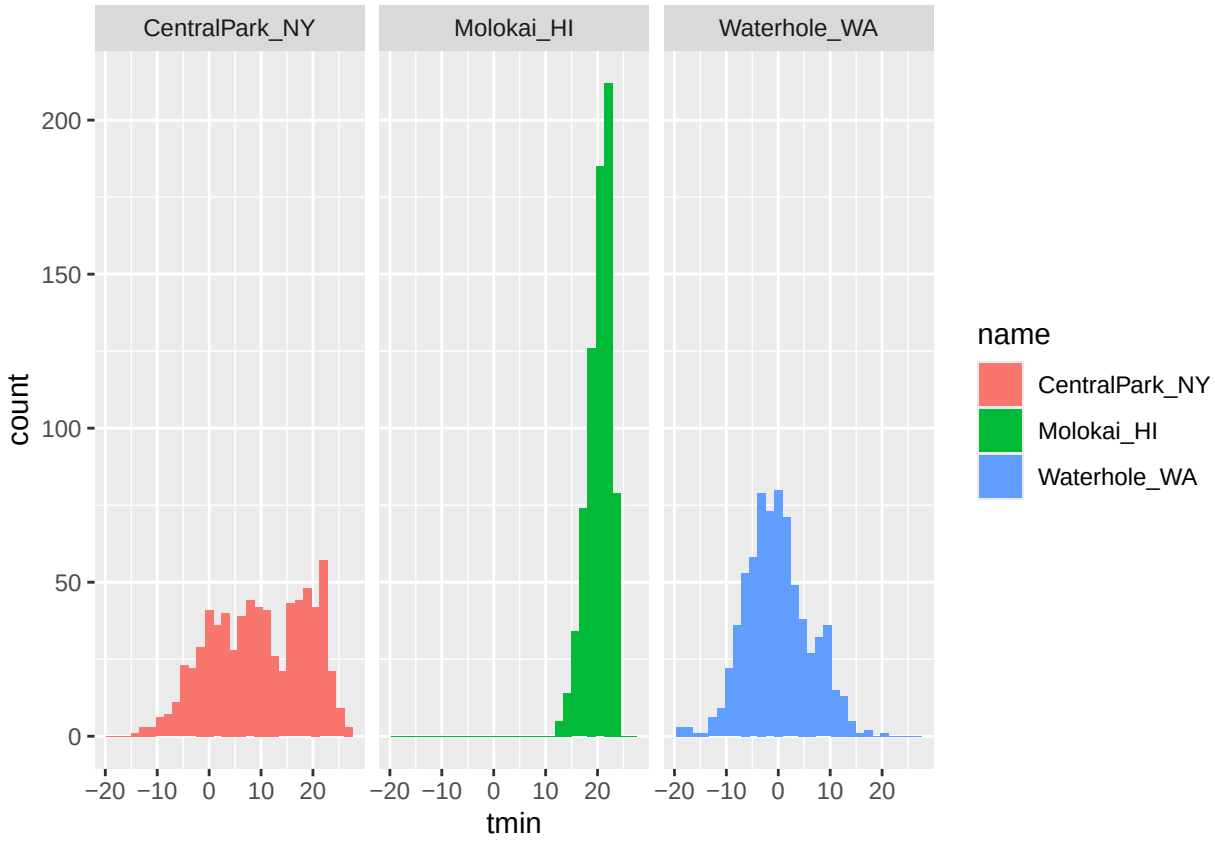


how would I fix this? maybe facet?

```
weather_df |>
  ggplot(aes( x = tmin, fill = name)) +
  geom_histogram() +
  facet_grid(. ~ name)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

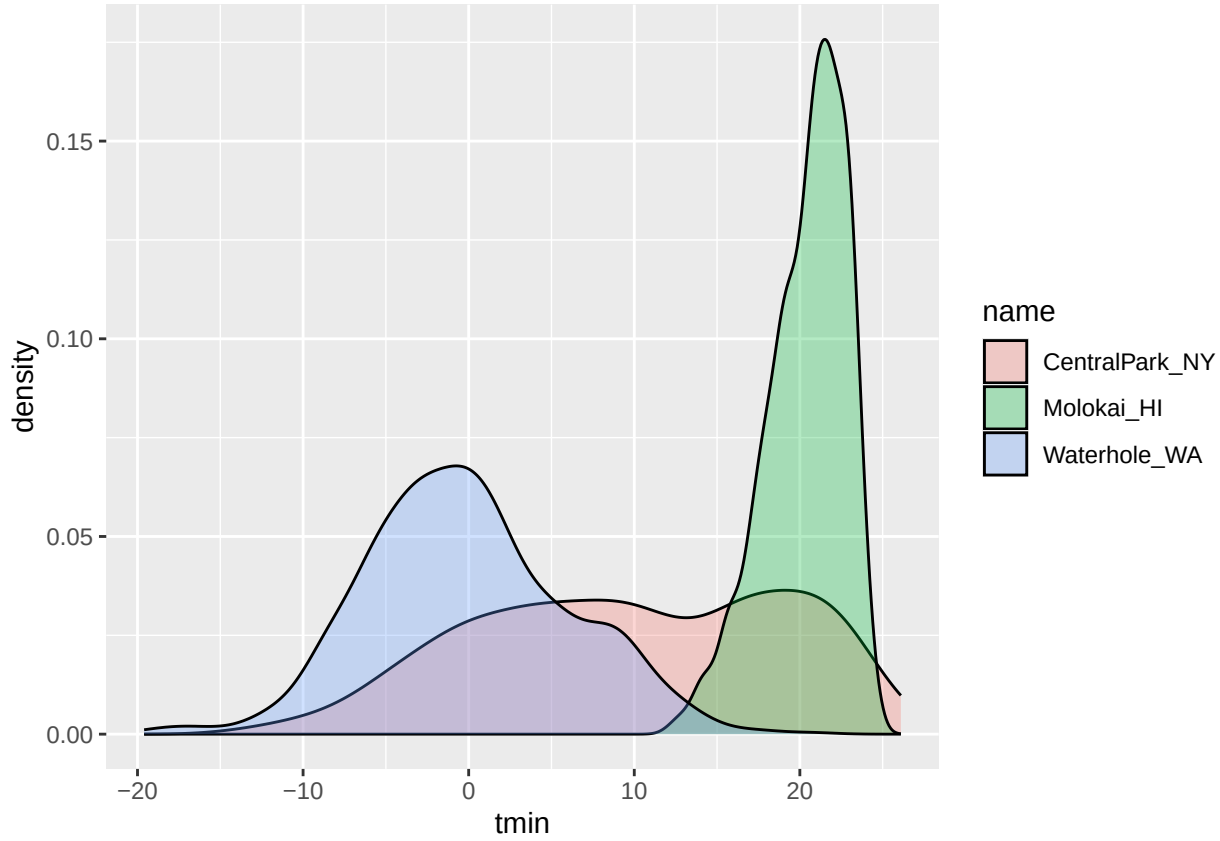
```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_bin()`).
```



maybe a density plot?

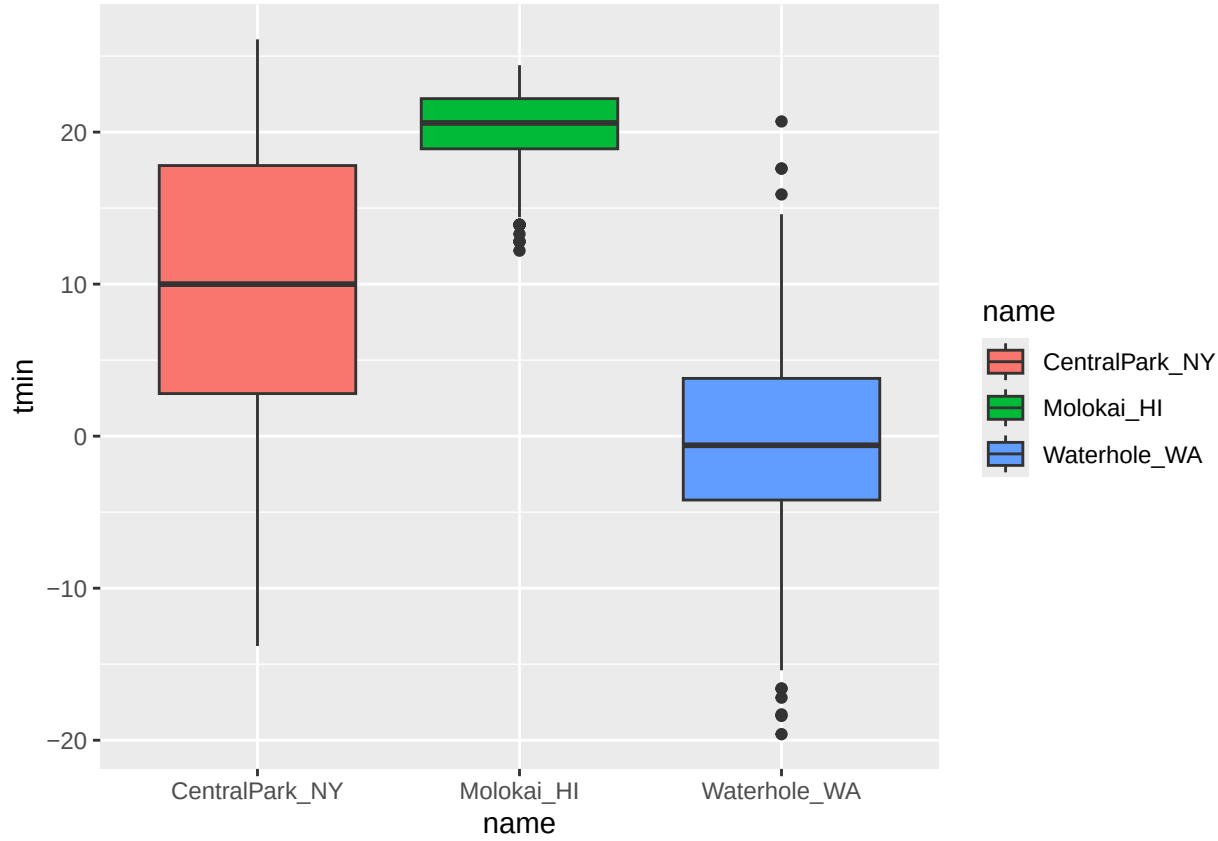
```
weather_df |>
  ggplot(aes(x = tmin, fill = name)) +
  geom_density(alpha = 0.3)
```

```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_density()`).
```




```
weather_df |>
  ggplot(aes(x = name, y = tmin, fill = name)) +
  geom_boxplot()
```

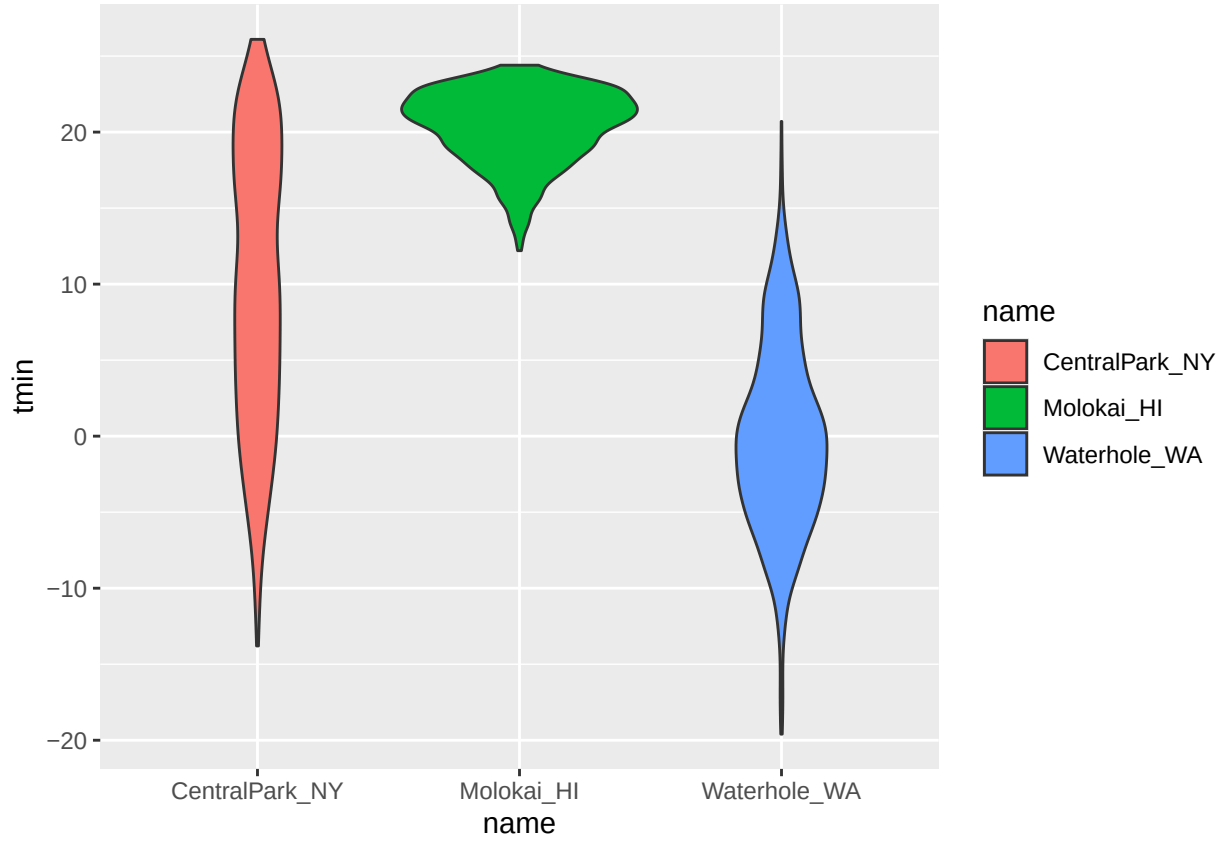
```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_boxplot()`).
```



violin plots

```
weather_df |>
  ggplot(aes(x = name, y = tmin, fill = name)) +
  geom_violin()
```

```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_ydensity()`).
```

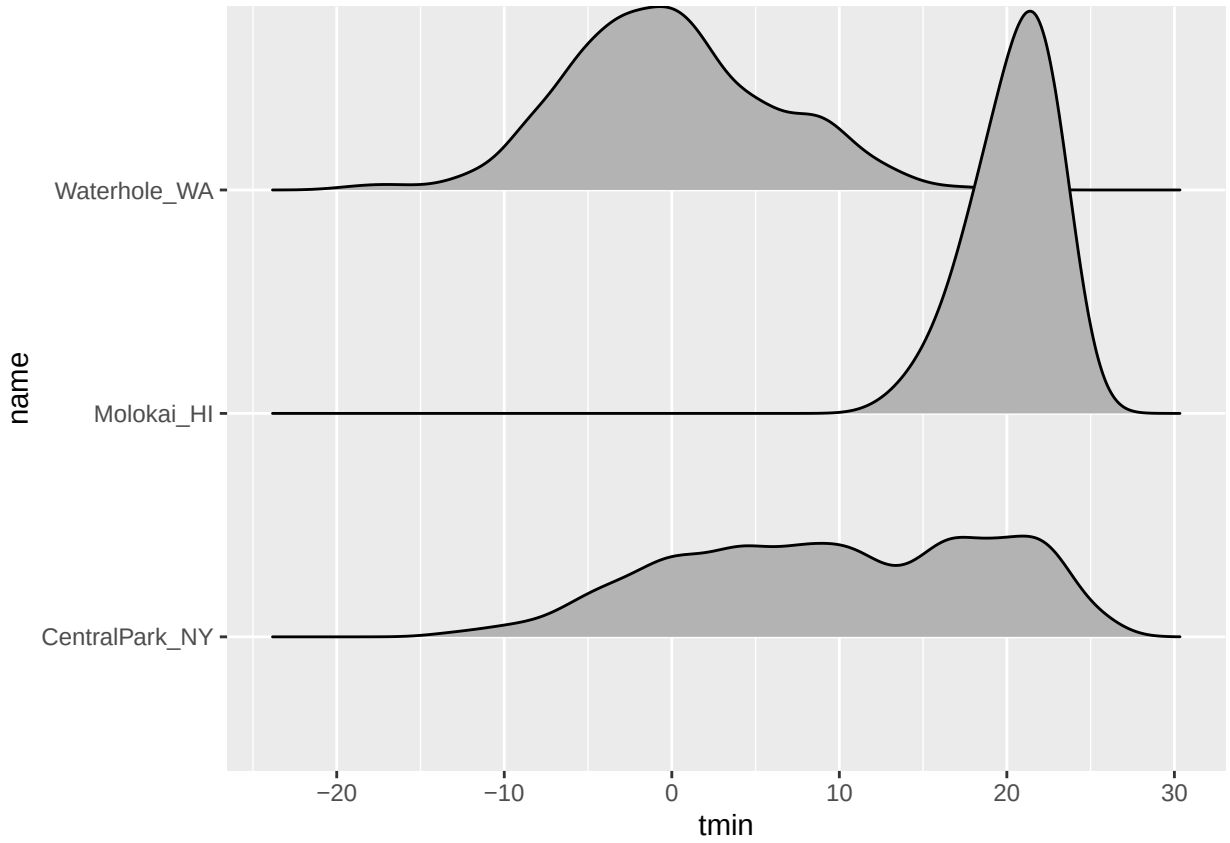


ridge plot

```
weather_df |>
  ggplot(aes(x = tmin, y = name)) +
  geom_density_ridges()
```

```
## Picking joint bandwidth of 1.41
```

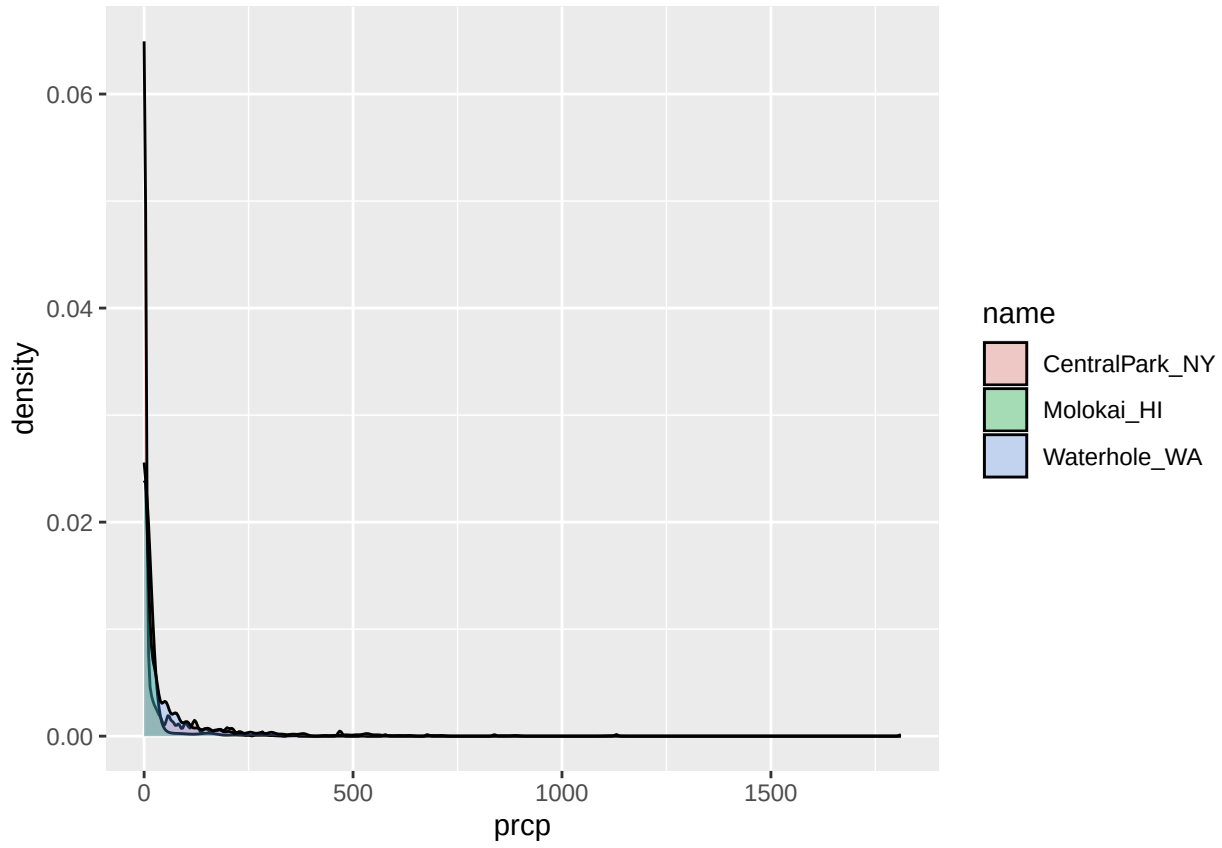
```
## Warning: Removed 17 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).
```



LA precipiration across locations

```
weather_df |>
  ggplot(aes(x = prcp, fill = name)) +
  geom_density(alpha = 0.3)
```

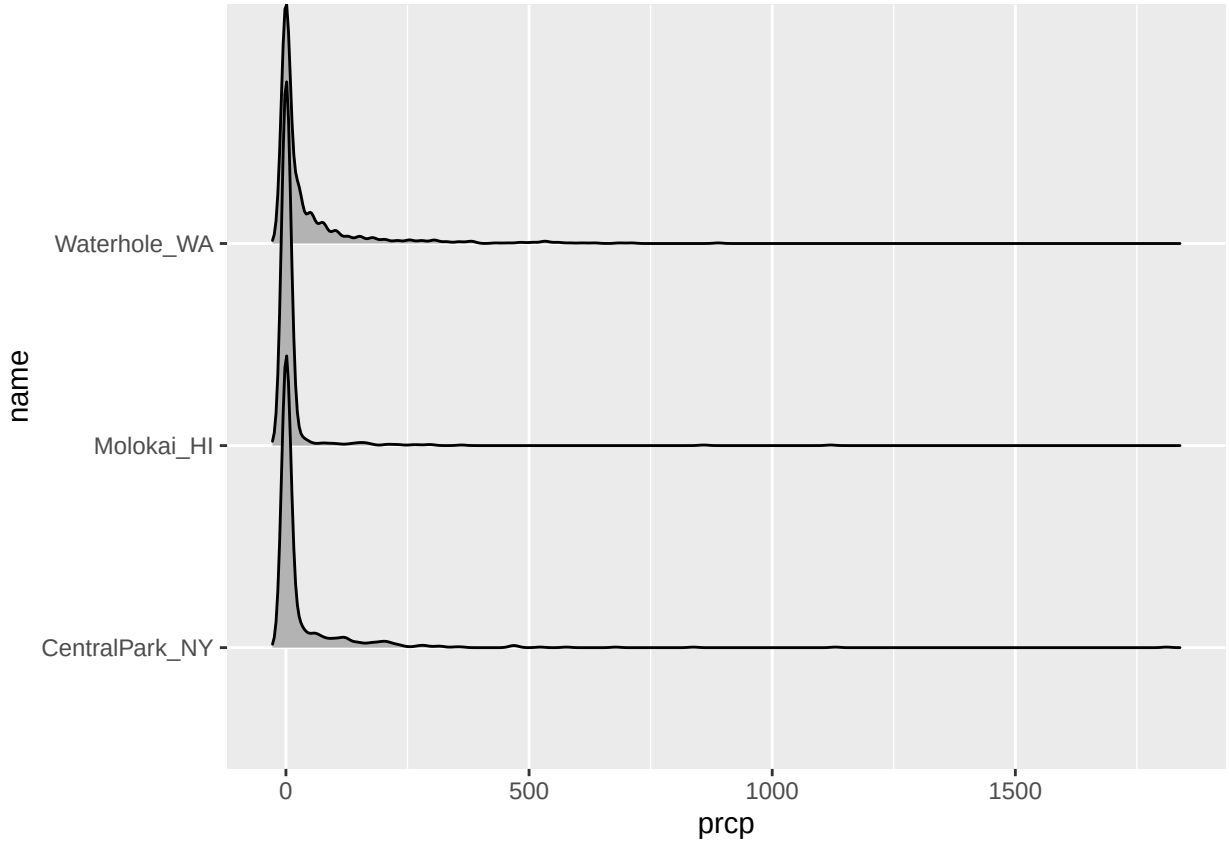
```
## Warning: Removed 15 rows containing non-finite outside the scale range
## (`stat_density()`).
```




```
weather_df |>
  ggplot(aes(x = prcp, y = name)) +
  geom_density_ridges()
```

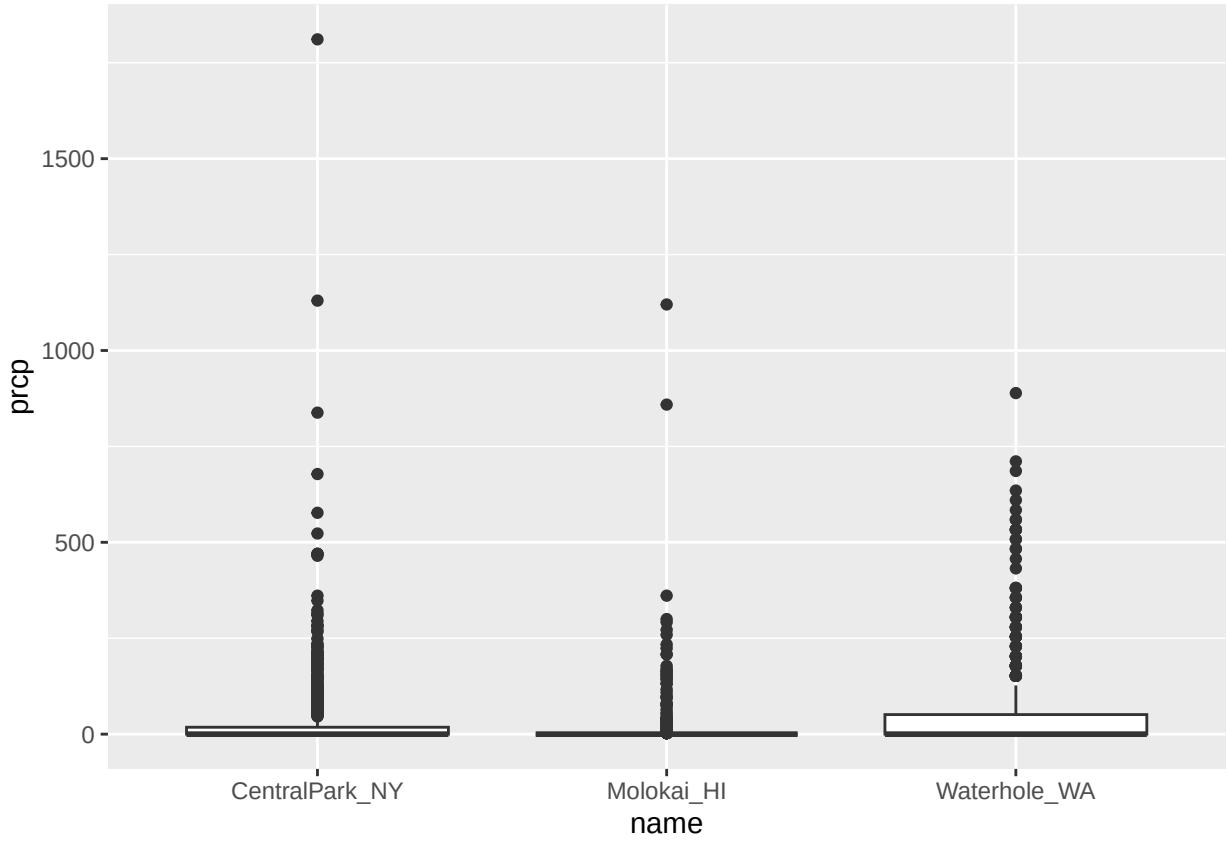
```
## Picking joint bandwidth of 9.22
```

```
## Warning: Removed 15 rows containing non-finite outside the scale range
## (`stat_density_ridges()`).
```




```
weather_df |>
  ggplot(aes(x = name, y = prcp)) +
  geom_boxplot()
```

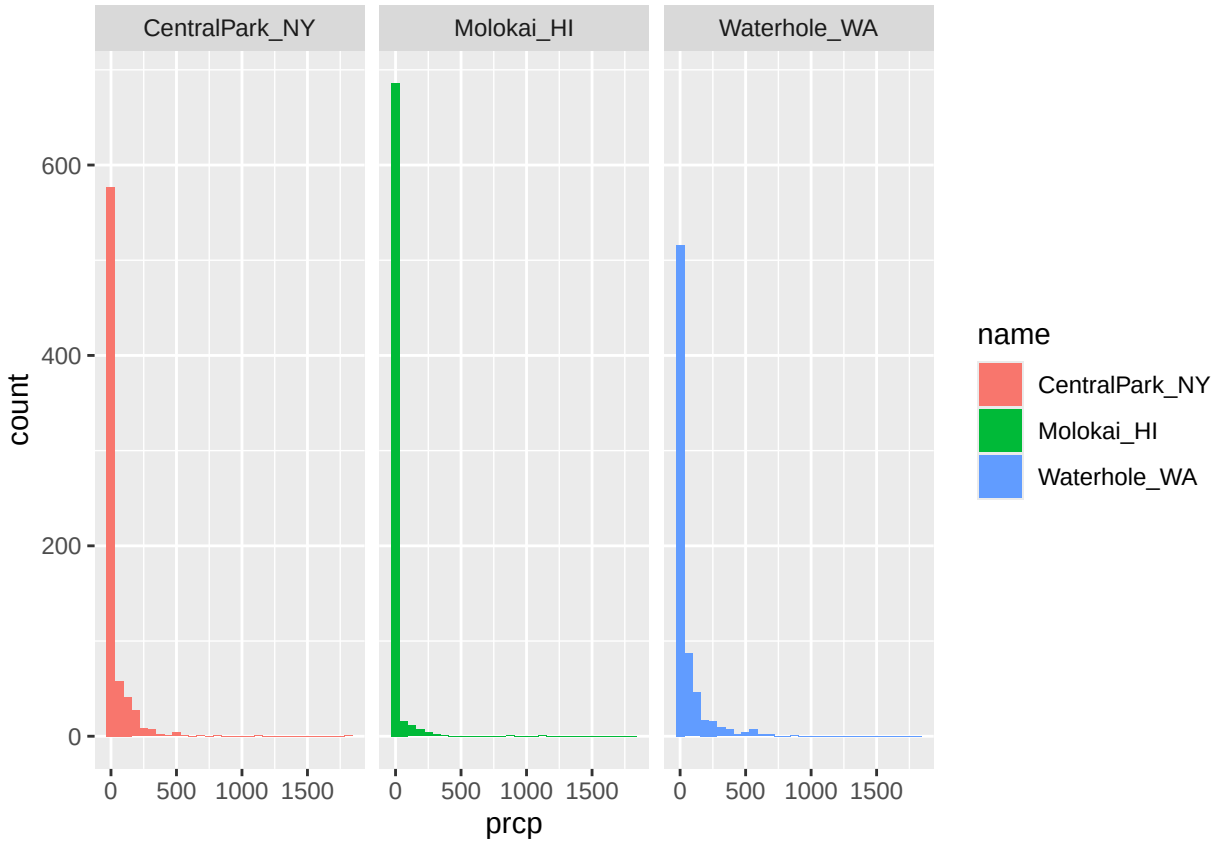
```
## Warning: Removed 15 rows containing non-finite outside the scale range
## (`stat_boxplot()`).
```



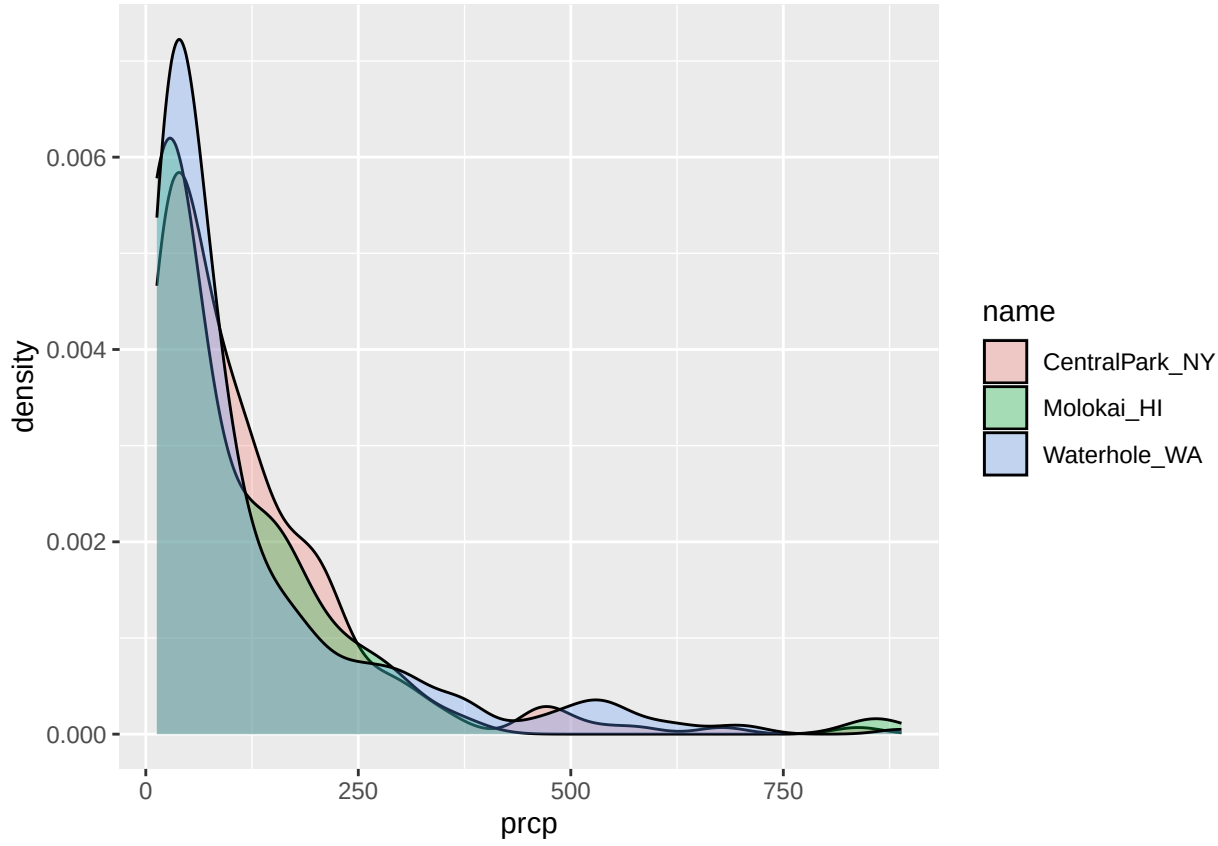

```
weather_df |>
  ggplot(aes(x = prcp, fill = name)) +
  geom_histogram() +
  facet_grid(. ~ name)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 15 rows containing non-finite outside the scale range
## (`stat_bin()`).
```




```
weather_df |>
  filter(prcp > 10, prcp < 1000) |>
  ggplot(aes(x = prcp, fill = name)) +
  geom_density(alpha = 0.3)
```



Saving and embedding plots

saving plots

```
ggp_weather =
  weather_df |>
  ggplot(aes(x = date, y = tmax, color = name)) +
  geom_point()

ggsave("plots/ggp_weather.pdf", ggp_weather, width = 8, height = 6)
```

```
## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

embedding plots

```
weather_df |>
  ggplot(aes(x = date, y = tmax, color = name)) +
  geom_point()
```

```
## Warning: Removed 17 rows containing missing values or values outside the scale range
## (`geom_point()`).
```

tmax
30
20
10
0
−10
2021−01
2021−07
2022−01
2022−07
2023−01
date
name
CentralPark_NY
Molokai_HI
Waterhole_WA