Class Activity 4

Your name here

2024-03-31

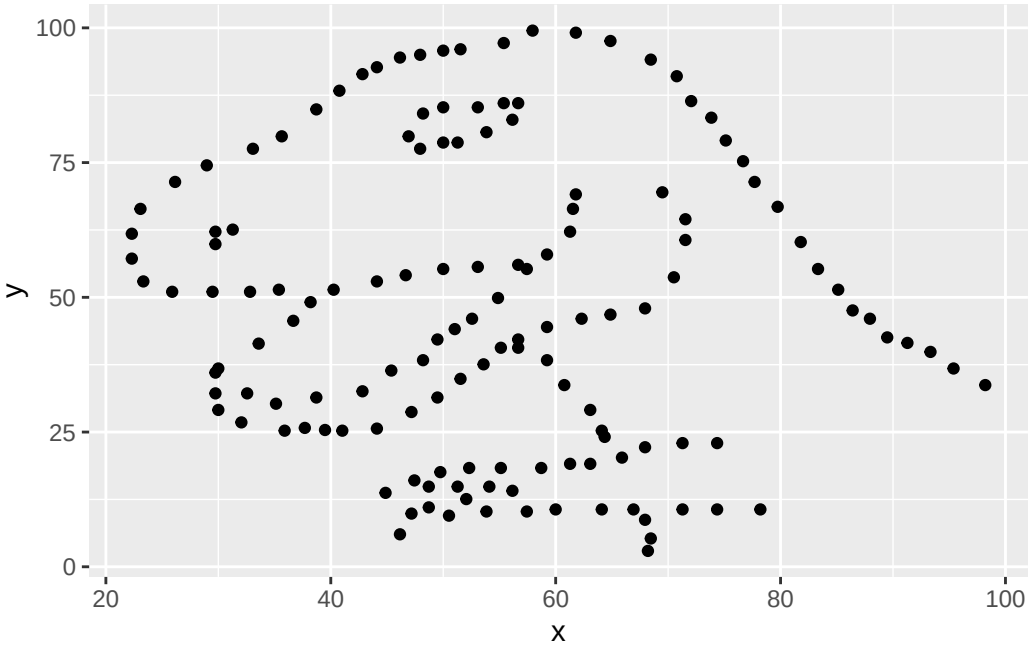
Your turn 1

This worksheet will guide you through creating various plots using the `ggplot2` package in R. We will be using the `datasaurus_dozen` dataset from the `datasauRus` package for demonstration purposes. The dataset contains 13 different datasets, and we'll use them to create a variety of plots.

Scatterplot

a. Run the following code.

```
ggplot(data = dino_data, mapping = aes(x = x, y = y)) +
  geom_point()
```



b. You *must* remember to put the aesthetic mappings in the `aes()` function! What happens if you forget?

Answer:

```
# Add a layer and see what happens
ggplot(data = dino_data , x = __ , y = __)
```

```
Error: <text>:2:32: unexpected input
1: # Add a layer and see what happens
2: ggplot(data = dino_data , x = __
                                  ^
```

c. The aesthetic mappings can be specified in the geom layer if you prefer, instead of the main `ggplot()` call. Give it a try:

Answer:

```
# Rebuild the scatterplot with your aesthetic mapping in the geom layer
ggplot(data = dino_data) +
   geom_point(____)
```

```
Error: <text>:3:16: unexpected input
2: ggplot(data = dino_data) +
3:    geom_point(__
                  ^
```

Bar Plot

In this problem, we'll explore creating a bar plot using the `datasaurus_dozen` dataset.

a. Create a new data frame called `dataset_counts` containing the count of observations in each dataset.

Answer:

```
_______ <- datasaurus_dozen %>%
  group_by(dataset) %>%
  summarise(count = n()) # number of rows in each dataset
```

```
Error: <text>:1:2: unexpected input
1: __
    ^
```

b. Create a bar plot showing the number of observations in each dataset.

Answer:

```
ggplot(data = _______, aes(x = dataset, y = count)) +
  geom_***(stat = "identity") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
Error: <text>:1:16: unexpected input
1: ggplot(data = __
                  ^
```

c. Generate a bar plot to visualize the median of the `x` variable across different datasets, with error bars denoting the interquartile range (IQR) for each dataset.

Answer:

```
# Calculate median and interquartile range for each dataset
dataset_summary <- datasaurus_dozen %>%
  group_by(dataset) %>%
  summarise(median_x = median(x), iqr_x = IQR(x))

# Create a bar plot with error bars representing the IQR
ggplot(dataset_summary, aes(x = , y = )) +
  geom_****() +
  geom_****() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(title = "Median of x by Dataset with IQR Error Bars", x = "Dataset", y = "Median of x
```

```
Error: <text>:8:10: unexpected '^'
7: ggplot(dataset_summary, aes(x = , y = )) +
8:   geom_****
            ^
```

Histogram

a. Create a histogram of the `x` variable for the `dino` dataset.

Answer:

```
ggplot(data = dino_data, aes(x = x)) +
  geom_****(binwidth = 2)
```

```
Error: <text>:2:10: unexpected '^'
1: ggplot(data = dino_data, aes(x = x)) +
2:   geom_****
            ^
```

b. Overlay a density curve on the histogram.

Answer:

```
ggplot(data = dino_data, aes(x = ____)) +
  geom_histogram(aes(y = ________), binwidth = ____, fill = ) +
  geom_density(color = "____")
```

```
Error: <text>:1:35: unexpected input
1: ggplot(data = dino_data, aes(x = __
                                    ^
```

Boxplot

Answer:

a. Create a boxplot of the x variable for each dataset in datasaurus_dozen.

```
ggplot(data = datasaurus_dozen, aes(x = dataset, y = x)) +
  geom_*****() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
Error: <text>:2:10: unexpected '^'
1: ggplot(data = datasaurus_dozen, aes(x = dataset, y = x)) +
2:   geom_****
            ^
```

Faceting

Answer:

a. Create a scatterplot of `x` vs. `y` for each dataset in `datasaurus_dozen` using `facet_wrap()`.

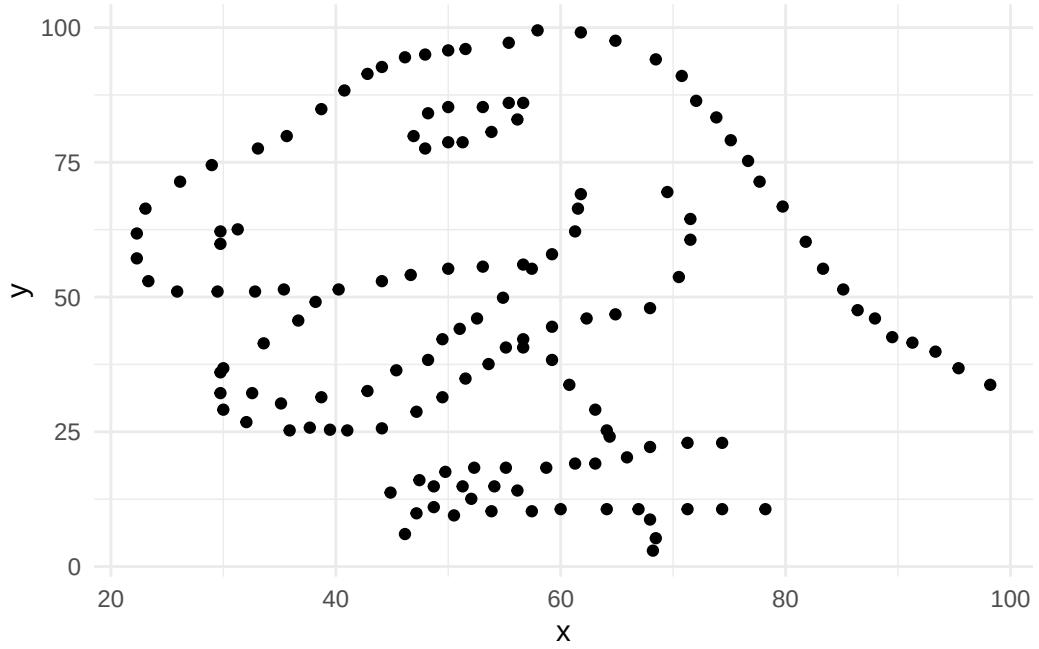
```
ggplot(data = datasaurus_dozen, aes(x = x, y = y)) +
  geom_point() +
  facet_***(~ ****) +
  theme_minimal()
```

```
Error: <text>:3:11: unexpected '*'
2:   geom_point() +
3:   facet_***
             ^
```

Variable Transformation

a. The scatterplot of the `dino` dataset without any transformations is given below. Complete the code and store the plot in an object called `p1`.

```
ggplot(data = dino_data, aes(x = x, y = y)) +
  geom_point() +
  theme_minimal()
```



b. Now, apply the square root transformation to both the `x` and `y` axes using the `scale_x_sqrt()` and `scale_y_sqrt()` functions in the `dino` dataset. Complete the code and store the plot in an object called `p2`.

Answer:

```
ggplot(data = dino_data, aes(x = x, y = y)) +
  geom_****() +
  scale_*****() +
  scale_*****() +
  theme_*****()
```

```
Error: <text>:2:10: unexpected '^'
1: ggplot(data = dino_data, aes(x = x, y = y)) +
2:   geom_****
            ^
```

c. Use `plot_layout()` and `plot_annotation` functions from `patchwork` package to plot the above two plots side-by-side.

```
library(patchwork)
# your r-code
```

d. Use `grid.arrange()` function from `gridExtra` package to get similar results as in part c.

Answer:

```
library(gridExtra)
# your r-code
```