Chapter 03: Individual geoms

Stan Piotrowski

September 20 2021

Contents

Introduction

This chapter focuses on exploring individual geoms: geoms which take `x` and `y` aesthetics, as well as color, size, and fill.

Exercises

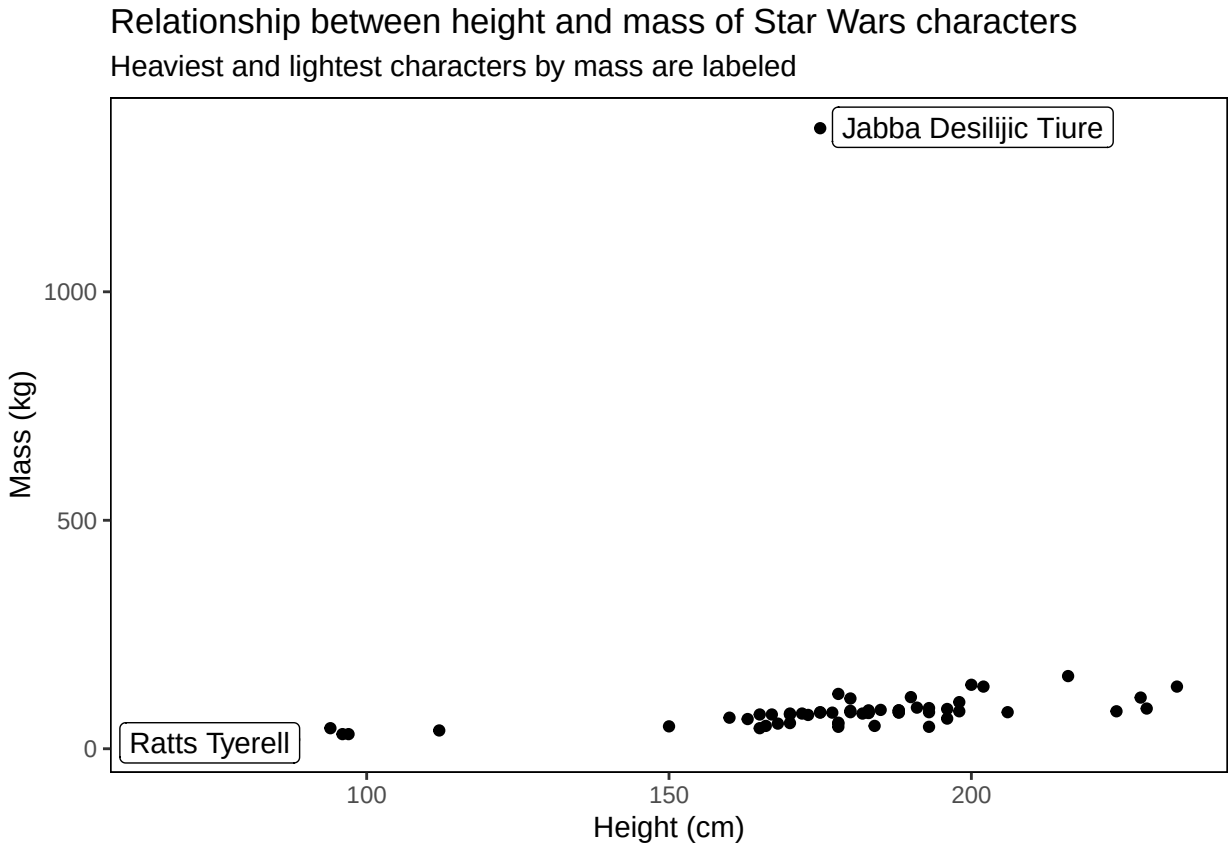
1) *What geoms would you use to draw each of the following named plots?*

For each of the following plots, we'll use the `starwars` dataset from dplyr, which contains data on different characters in the saga including name, height, mass, home planet, the films they appear in, species, and other information.

```
## Rows: 87
## Columns: 14
## $ name       <chr> "Luke Skywalker", "C-3PO", "R2-D2", "Darth Vader", "Leia Or~
## $ height     <int> 172, 167, 96, 202, 150, 178, 165, 97, 183, 182, 188, 180, 2~
## $ mass       <dbl> 77.0, 75.0, 32.0, 136.0, 49.0, 120.0, 75.0, 32.0, 84.0, 77.~
## $ hair_color <chr> "blond", NA, NA, "none", "brown", "brown, grey", "brown", N~
## $ skin_color <chr> "fair", "gold", "white, blue", "white", "light", "light", "~
## $ eye_color  <chr> "blue", "yellow", "red", "yellow", "brown", "blue", "blue",~
## $ birth_year <dbl> 19.0, 112.0, 33.0, 41.9, 19.0, 52.0, 47.0, NA, 24.0, 57.0, ~
## $ sex        <chr> "male", "none", "none", "male", "female", "male", "female",~
## $ gender     <chr> "masculine", "masculine", "masculine", "masculine", "femini~
## $ homeworld  <chr> "Tatooine", "Tatooine", "Naboo", "Tatooine", "Alderaan", "T~
## $ species    <chr> "Human", "Droid", "Droid", "Human", "Human", "Human", "Huma~
## $ films      <list> <"The Empire Strikes Back", "Revenge of the Sith", "Return~
## $ vehicles   <list> <"Snowspeeder", "Imperial Speeder Bike">, <>, <>, <>, "Imp~
## $ starships  <list> <"X-wing", "Imperial shuttle">, <>, <>, "TIE Advanced x1",~
```

1.1) _Scatterplot_

For a scatterplot, we would use the `goem_point()` geom. Let's use a scatterplot to explore the relationship between the height (cm) and mass (kg) of different Star Wars characters. Note, we'll remove any entries with missing data in either of these fields before plotting.

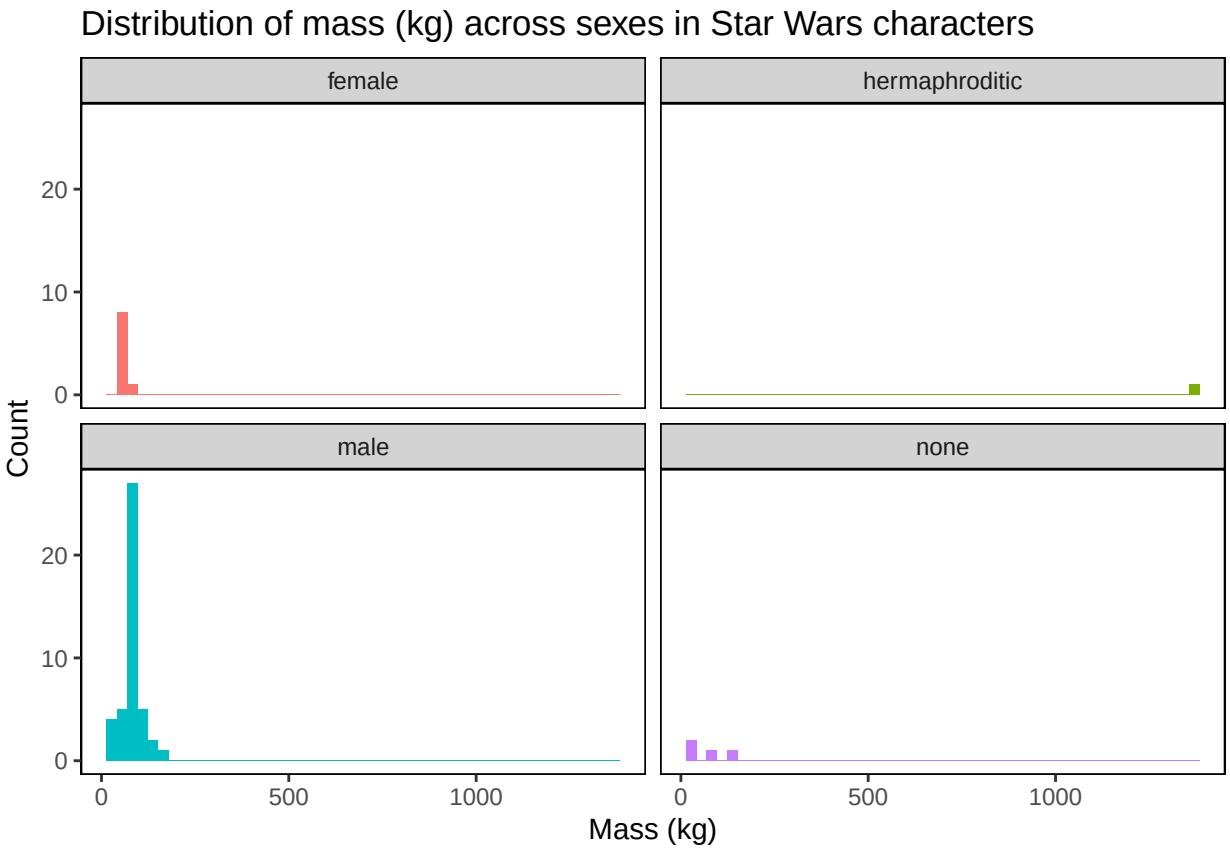


`1.2) _Line chart_`

To create a line chart to connect observations in the order of the variable on the x-axis, we would use the `geom_line()` geom.

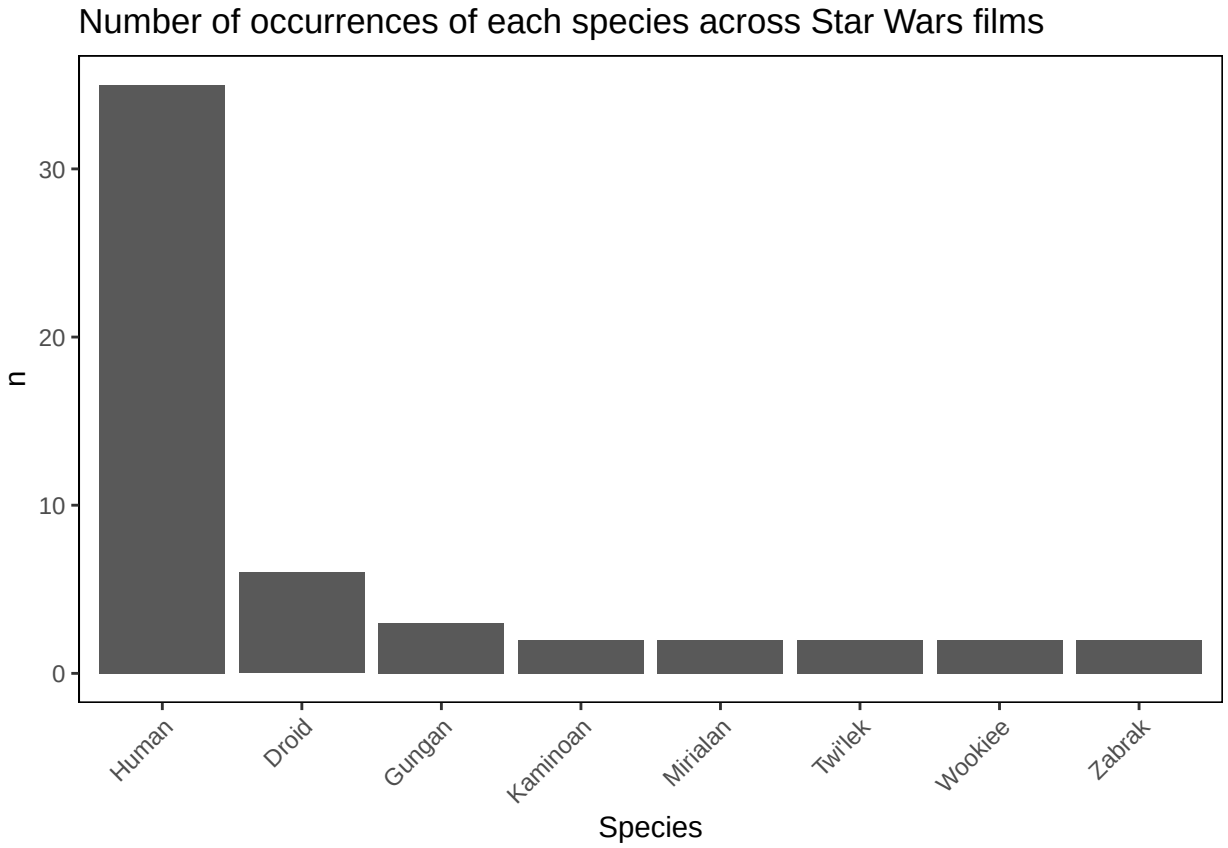
`1.3) _Histogram_`

To create a histogram, we would use the `geom_histogram()` geom. Alternatively, we could use the `geom_bar()` geom and use a statistical transformation to bin and count observations. For example, we can visualize the distribution of mass across sexes.



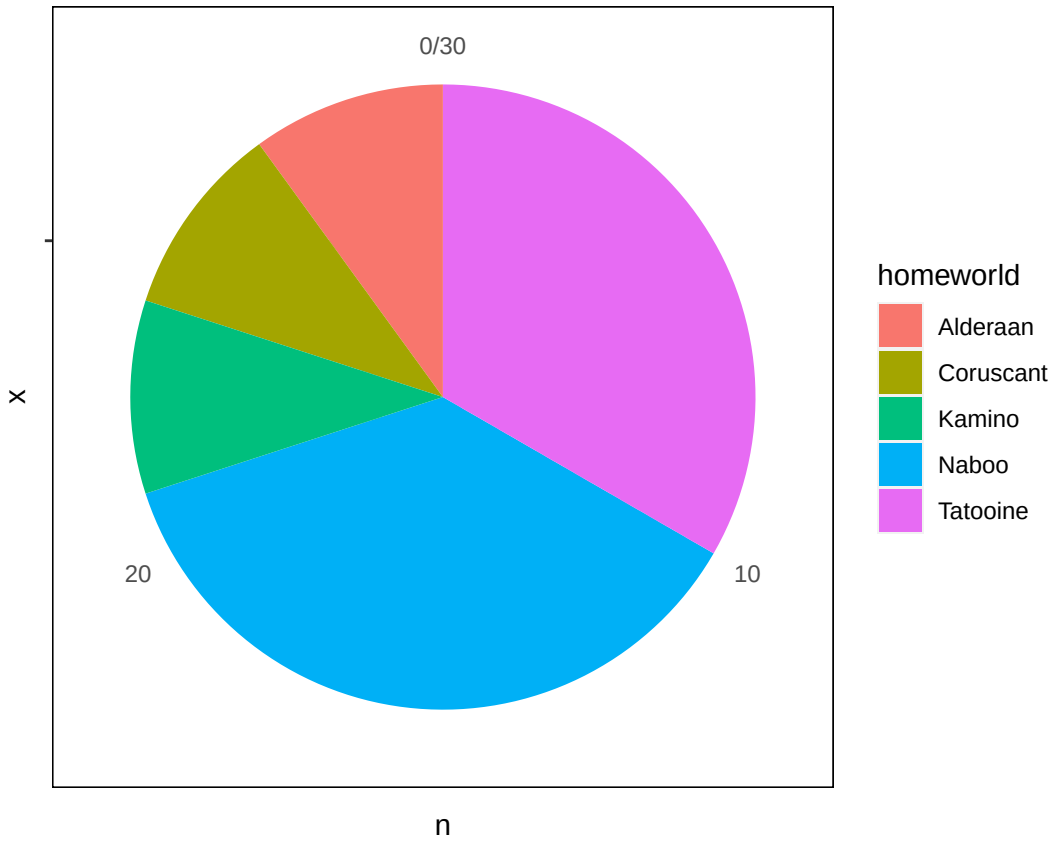
1.4) _Bar chart_

To create a bar chart, we would use the `geom_bar()` geom with the `stat = "identity"` argument to ensure that the data are plotted "as is" without statistical transformation. We can create a bar chart of the number of different characters by species overall.



`1.5) _Pie chart_`

To draw a pie chart, we'll use the `geom_bar()` geom with `stat = "identity"` argument and use the `coord_polar` coordinate system. We'll use a pie chart to plot the proportion of characters from each home world, subsetted by the home worlds with the highest number of characters. A more thorough description for creating pie charts with ggplot2 can be found here: https://www.r-graph-gallery.com/piechart-ggplot2.html.



2) *What's the difference between* `geom_path()` *and* `geom_polygon()`*? What's the difference between* `geom_path()` *and* `geom_tile()`*?*

3) *What low-level geoms are used to draw* `geom_smooth()`*? What about* `geom_boxplot()` *and* `geom_violin()`*?*