Context

Getting into the Pirate business is rough. Ending up with the wrong crewmates is sure to be a disaster for this career and can likely lead to having to "walk the plank". Thankfully, Nathaniel D. Phillips did some good work with curating and providing a survey of 1000 Pirates from the 2015 annual international pirate meeting at the Bodensee in Konstanz, Germany which can help aspiring pirate captains choose their crewmates wisely.

Using the `dplyr` package's `glmpse()` function we can get a quick look at all the surveyed characteristic of the `pirates` data set (for a full description of the data be sure to write `?pirates` in the console)

```
# Upload the necessary libraries
library(yarrr)
library(tidyverse)
library(ggthemes)
glimpse(pirates)
```

```
## Rows: 1,000
## Columns: 17
## $ id               1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,
16,...
## $ sex              "male", "male", "male", "female", "female", "male",
"f...
## $ age              28, 31, 26, 31, 41, 26, 31, 31, 28, 30, 25, 20, 24,
26...
## $ height           173.11, 209.25, 169.95, 144.29, 157.85, 190.20,
158.05...
## $ weight           70.5, 105.6, 77.1, 58.5, 58.4, 85.4, 59.6, 74.5,
68.7,...
## $ headband         "yes", "yes", "yes", "no", "yes", "yes", "yes",
"yes",...
## $ college          "JSSFP", "JSSFP", "CCCC", "JSSFP", "JSSFP", "CCCC",
"J...
## $ tattoos          9, 9, 10, 2, 9, 7, 9, 5, 12, 12, 10, 14, 8, 9, 14, 8,
...
## $ tchests          0, 11, 10, 0, 6, 19, 1, 13, 37, 69, 1, 5, 6, 12, 70,
3...
## $ parrots          0, 0, 1, 2, 4, 0, 7, 7, 2, 4, 3, 3, 0, 3, 0, 1, 0, 3,
...
## $ favorite.pirate  "Jack Sparrow", "Jack Sparrow", "Jack Sparrow", "Jack
...
## $ sword.type       "cutlass", "cutlass", "cutlass", "scimitar",
"cutlass"...
## $ eyepatch         1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0,
...
## $ sword.time       0.58, 1.11, 1.44, 36.11, 0.11, 0.59, 3.01, 0.06,
0.74,...
## $ beard.length     16, 21, 19, 2, 0, 17, 1, 1, 1, 25, 1, 27, 0, 19, 0,
1,...
```

```
## $ fav.pixar      "Monsters, Inc.", "WALL-E", "Inside Out", "Inside
Out"...
## $ grogg          11, 9, 7, 9, 14, 7, 9, 12, 16, 9, 7, 8, 12, 7, 9, 10,
...
```

There's alot of characteristics to be looking for for an ideal crew. For the sake of berevity, in this blog we will look at sword drawing speed. (Hopefully, in future blogs we will look at other characteristics)

So without further ado, lets Explore data with R!(Pirate voice intended)

Finding the Best Sword Draw- Looking at Education and Swordtype

Because we are just interested in Pirates who posess their own swords (and not a banana). Lets see fror whom we can get the best sword draw- and which college to recruit them from.

Our data set has two pirate schools – Captian Chunk's Cannon Crew (CCCC) and Jack Sparro's School of Fashion and Piratery.

```
ggplot(
  data = filter(pirates, sword.type != "banana"),
  mapping = aes(x = sword.type, y = sword.time)
) +
  geom_violin(mapping = aes(fill = sword.type), show.legend = F) +
  facet_wrap( ~ college)+
  ggtitle("Sword Draw Speed vs Sword Type (by College)")+
  theme_solarized_2()+
  theme(plot.title = element_text(hjust = 0.5))
```
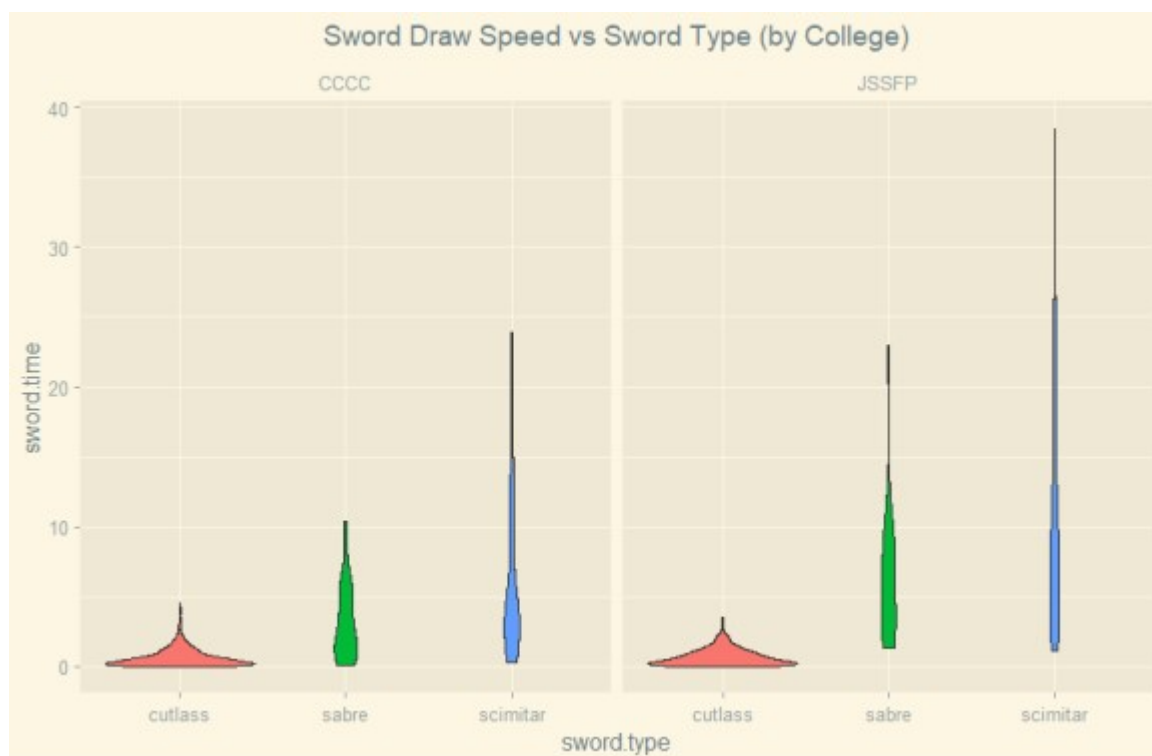

From the voilin charts it seems that recruiting cutlass weilders from either Jack Sparro's or Captian Chunk's should be fine. But for sabre and scimitar specialists, our best candidates are from Captian Chunk's Cannon Crew.

Assuming that our prospective crew members will submit a resume to join the crew, this is a great way to recruit swordsmen for our crew. But usually with being a pirate, having a resume is not readily available and asking these questions to every single candidate can be quite time consuming.

What if we don't know about their schooling or about the weapons that they use? What can we use as a heuristic? Well, we can do what pirates usually do. Size a matey up!

Finding the Best Sword Draw by "Sizin' 'em up".

Besides for looking at a Resume, it is probably a good idea to have some heuristic with how to select an ideal pirates. There's only so much which can be said on paper, but in a moment of crisis- how can we be sure if they will stand up to the challenge and be able to swashbuckle with the best of 'em?

Choosing mateys is no laughing matter, lets see what we can offer as advice to our Pirate recruiters!

Height and Weight

Looking at the plot between hight and weight we see that there is a very high correlation between the two. With this in mind we can use either of them to be one of our "sizing up" criterion.

```
ggplot(data = pirates, mapping = aes(x = weight, y = height)) +
  geom_point() +
  geom_smooth(formula = y ~ x)+
  ggtitle("Pirate Height vs Weight")+
  theme_solarized_2()+
  theme(plot.title = element_text(hjust = 0.5))
```



```
# correlation between weight and height
cor(pirates$height, pirates$weight)
```

```
## [1] 0.9318938
```

Because height and weight have a very high correlation, we can use the one which will account for the other to some degree. Because it is easier to estimate height, we will go with that.

Bandanas, Eyepatches and Tattoos.

Looking at our filtered data plotting height against sword time and faceting on whether or not the Pirate has a bandana, an eyepatch and coloring the points based on number of tatoos a pirate has (assuming they are all visible) we see the following.

- While height doesn't seem to be inicative of good sword-draw speed, a potential crewmate showing up with a headband or an eyepatch does.
- A pirate with a bandana appears to have better sword time than a pirate without one.
- A pirate with a eyepatch has poorer swordtime than a pirate without it. This is probably because seeing with just one eye results in having poorer depth perception((?) Just a thought.)

```
ggplot(data = filter(pirates, sword.type != "banana")) +
  geom_point(mapping = aes(x = height, y = sword.time, color = tattoos)) +
  facet_grid(eyepatch ~ headband) +
  scale_color_gradient2(
    midpoint = mean(pirates$tattoos),
    low = "orange",
    mid = "blue",
    high = "black"
  )+
  ggtitle("Pirate Height vs Sword Time (with Headbands/Eypatch comparison)")+
  theme_solarized_2()+
  theme(plot.title = element_text(hjust = 0.5))
```
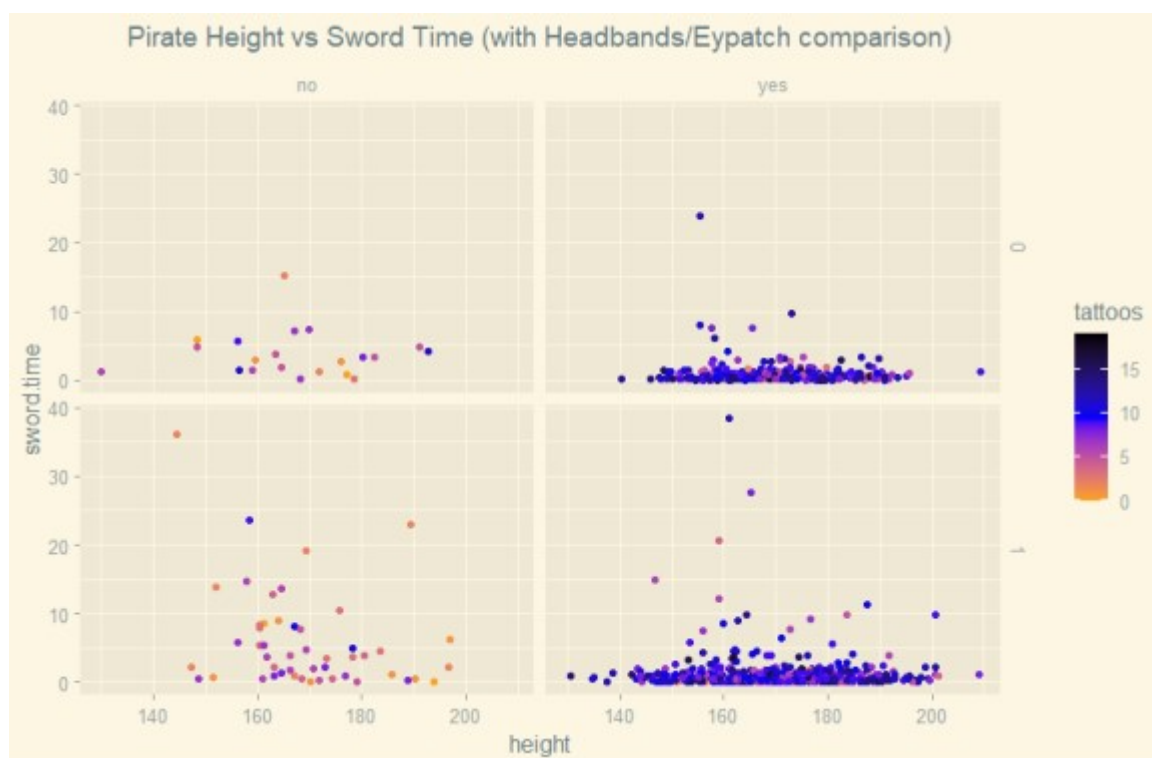

With this plot here, we can't say anything conclusive about Tattoos being a good indicator of good swordmanship.

Lets focus on the relationship between number of Tattoos and sword time. We will also facet on whether a

given pirate has a headband or an eyepatch.

```
ggplot(data = filter(pirates, sword.type != "banana")) +
  geom_point(
    mapping = aes(x = tattoos, y = sword.time, color = tattoos),
    show.legend = FALSE
  ) +
  facet_grid(eyepatch ~ headband) +
  scale_color_gradient2(
    midpoint = mean(pirates$tattoos),
    low = "orange",
    mid = "blue",
    high = "black"
  ) +
  ggtitle("Tattoo number vs Sword Time (with Headbands/Eypatch comparison)")
  theme_solarized_2() +
  theme(plot.title = element_text(hjust = 0.5))
```
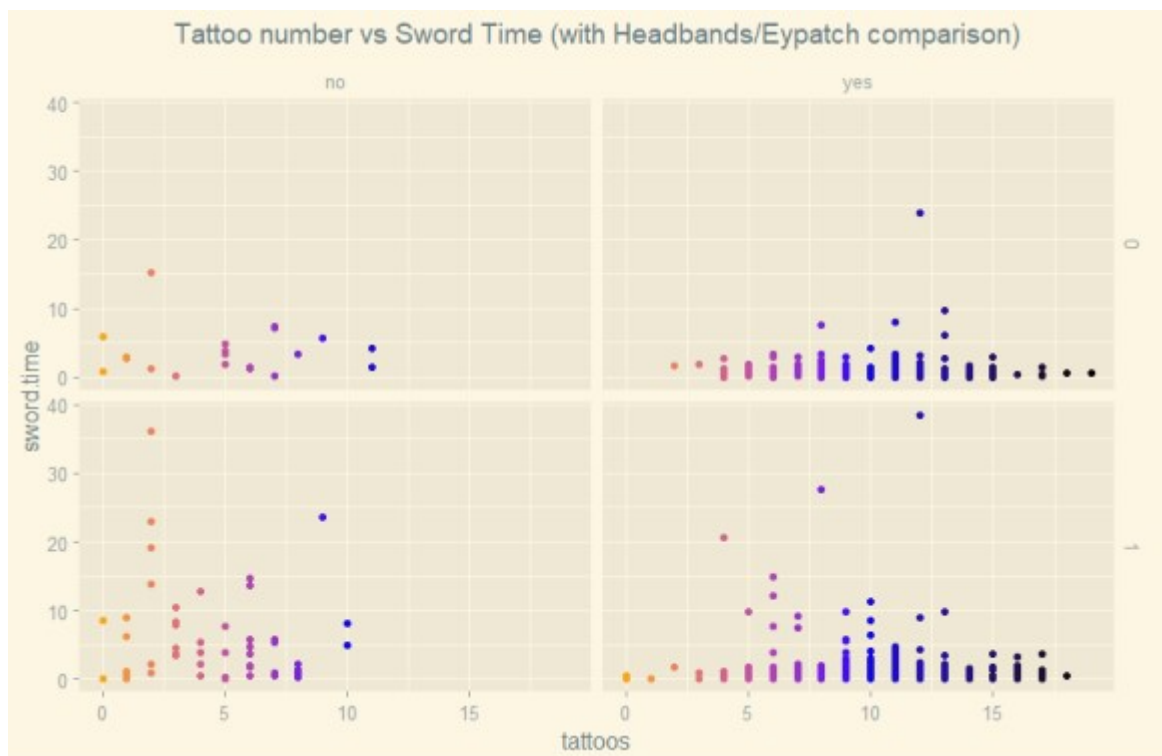

From this plot we see the following:

1. Pirates with a bandana have on average (this was done by eye-balling [no pun intended]) a better sword draw than pirates without one.
2. Pirates with over 14 tattoos are likely to possess a superior sword draw. However, those without eyepatches tend to do better than those with.
3. If you are going to recruit pirates with eyepatches (which you likely will want if you're running a pirate ship) – you will want to be sure they have many tattoos to assure that they have a good draw

While there are exceptions to the rule, these are some good hueristics to follow and recommend that othe Pirate recruiters have these in mind when looking for ideal swordsmen for their crew.

Conclusion

This was a really fun dataset to explore. I hope to do more exploring in future blogs on this topic where we're going to try to get some heuristics for finding Pirates who will help us find treasures and who is idea for rationing grogg while sailing the seven seas!