

Data exploration

Anonymous

2021-10-18

1. Prepare the data

```
"C:\\Users\\maxim\\UQAM\\Montiglio, Pierre-Olivier - Data Behaviour"
```

1.1 Load the required packages

```
using Pkg, Weave, CSV, DataFrames, Plots
```

1.2 Import the data

```
data = DataFrame(CSV.File(joinpath(path, "03_final-data",  
                                     "03_final-data_2021",  
                                     "final-data.csv")))
```

2. Inspect the data

2.1 Distribution of predator behaviour

In this section, we will inspect the distribution of the different predator behaviors. We provide the distribution of the raw variables and their transformations.

The transformations include :

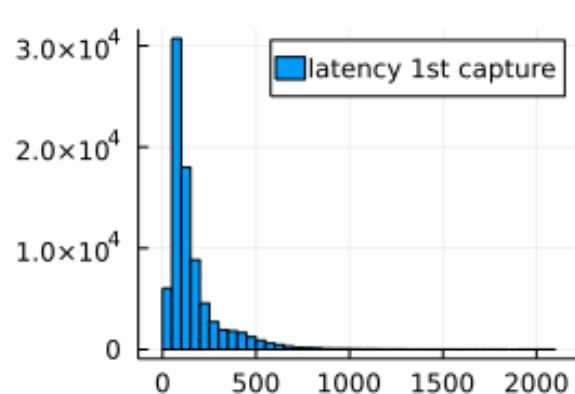
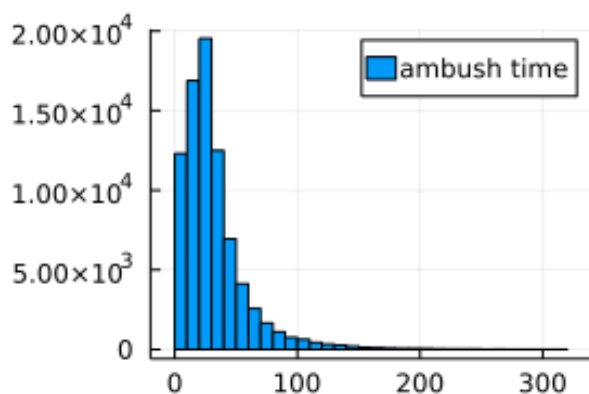
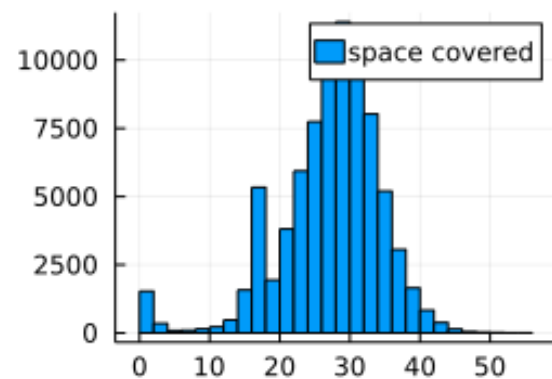
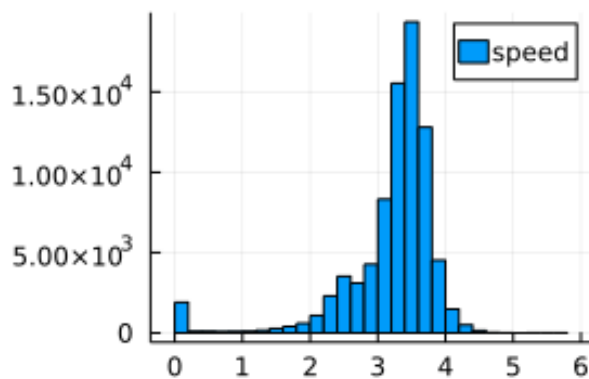
- square root
- log

2.1.1 Raw variables

```
# Predator behavior
a = histogram(data.pred_speed,
              bins = 40,
              label = "speed")
b = histogram(data.pred_amount_tiles_visited,
              bins = 40,
              label = "space covered")
c = histogram(data.ambush_time_close,
              bins = 40,
              label = "ambush time")
d = histogram(data.latency_1st_capture,
              bins = 40,
              label = "latency 1st capture")

e = histogram(data.total_chase_duration,
              bins = 40,
              label = "chase duration")

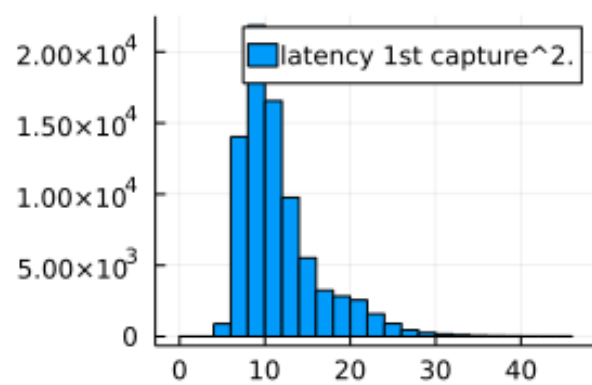
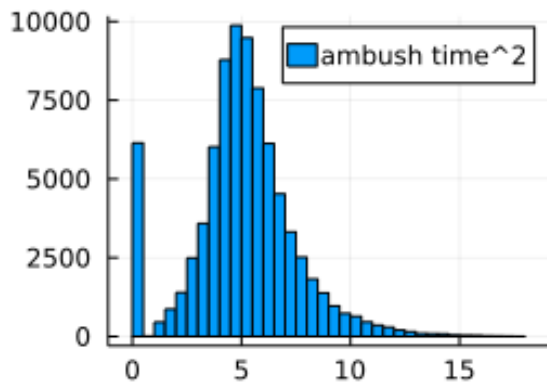
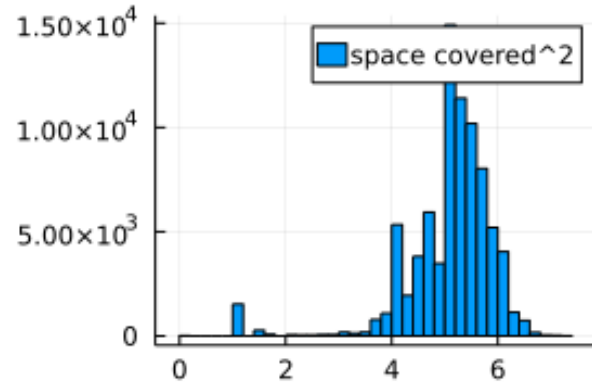
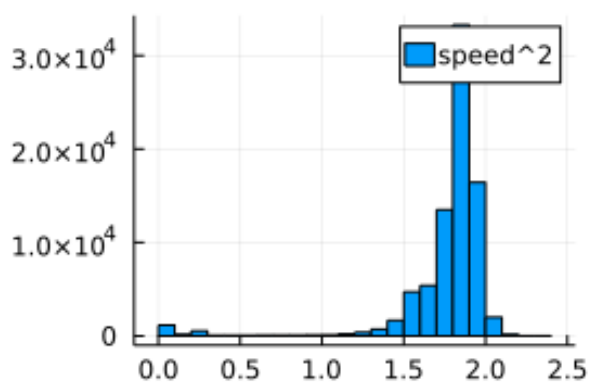
plot(a, b, c, d, layout = 4)
```



2.1.2 Square root variables

```
# Square root of predator behavior
a1 = histogram(sqrt.(data.pred_speed),
               bins = 40,
               label = "speed^2")
b1 = histogram(sqrt.(data.pred_amount_tiles_visited),
               bins = 40,
               label = "space covered^2")
c1 = histogram(sqrt.(data.ambush_time_close),
               bins = 40,
               label = "ambush time^2")
d1 = histogram(sqrt.(data.latency_1st_capture),
               bins = 40,
               label = "latency 1st capture^2.")

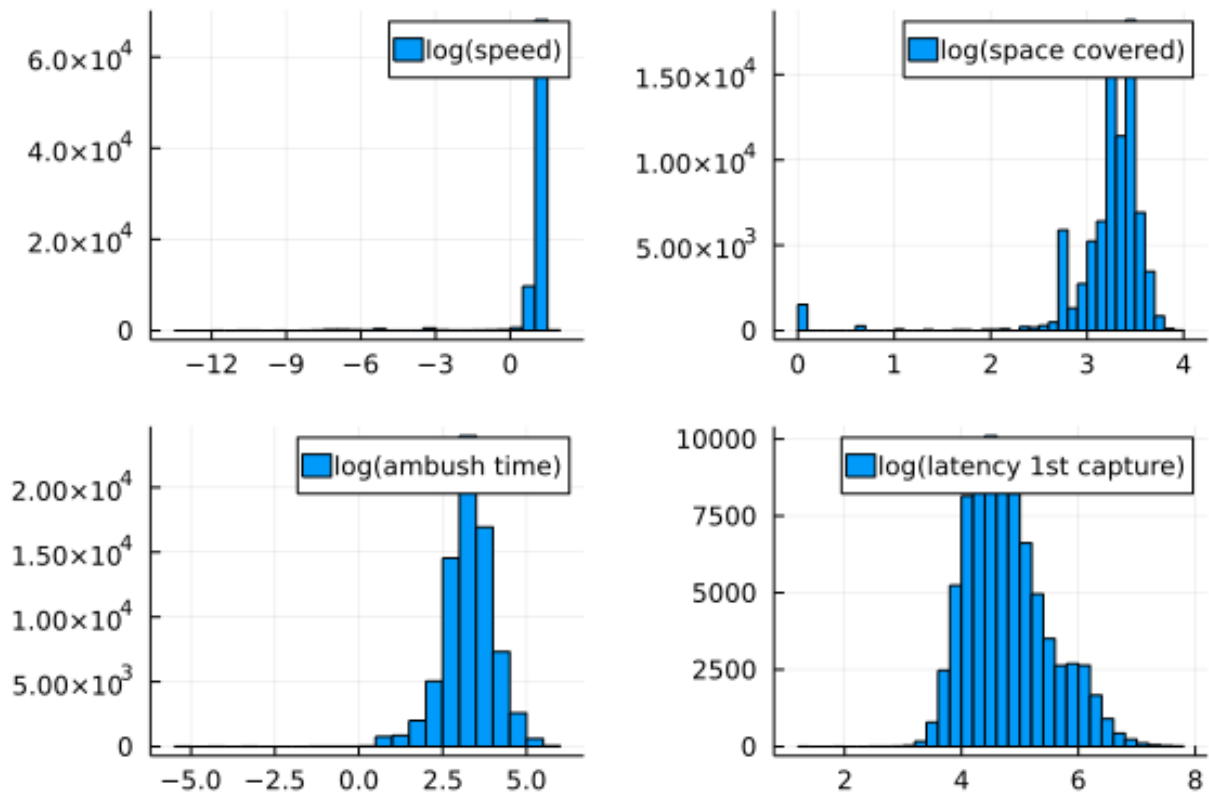
plot(a1, b1, c1, d1, layout = 4)
```



2.1.3 Log variables

```
# Log of predator behavior
a2 = histogram(log.(data.pred_speed),
               bins = 40,
               label = "log(speed)")
b2 = histogram(log.(data.pred_amount_tiles_visited),
               bins = 40,
               label = "log(space covered)")
c2 = histogram(log.(data.ambush_time_close),
               bins = 40,
               label = "log(ambush time)")
d2 = histogram(log.(data.latency_1st_capture), bins = 40,
               label = "log(latency 1st capture)")

plot(a2, b2, c2, d2, layout = 4)
```



2.1 Distribution of prey behavior

We now inspect the distribution of the prey behavior. The different prey behaviors were calculated using the average behavior of the 4 prey within a match. We provide the distribution of the raw variables and their transformations.

The transformations include :

- square root
- log

Prey behavior

```
f = histogram(data.prey_avg_speed,
              bins = 40,
              label = "prey speed")
g = histogram(data.prey_avg_amount_tiles_visited,
              bins = 40,
              label = "prey space covered")

h = histogram(data.prey_total_heal_count,
              bins = 40,
              label = "prey heal count")

i = histogram(data.prey_avg_amount_tiles_visited,
              bins = 40,
              label = "prey rescue count")
#g = histogram(data.boldness, bins = 50, label = "Prey boldness")
plot(f, g, h, i, layout = 4)
```

