



MACQUARIE
University
SYDNEY • AUSTRALIA

STAT 1378

Assignment 3: Investigation of top 500 passwords.

02 November 2021



- ▶ Investigation of top 500 passwords.
- ▶ 2 main questions:
- ▶ Which Category has the weakest passwords?
- ▶ Does having a stronger password affect how long it takes to crack it?
- ▶ Finals Conclusions.
- ▶ Other Recommendations.
- ▶ References

Variables:

1. **Rank** (Which is the most COMMON password)
2. **Password** (The password itself).
3. **Value** (Time to crack by online guessing)
4. **TimeUnit** (The metric of time for a password brute force.)
5. **OfflineCrackSec** (The amount of seconds for a password to be cracked)
6. **Strength** (Strength of password according to security.org)
7. **Font size**

Hypothesis:

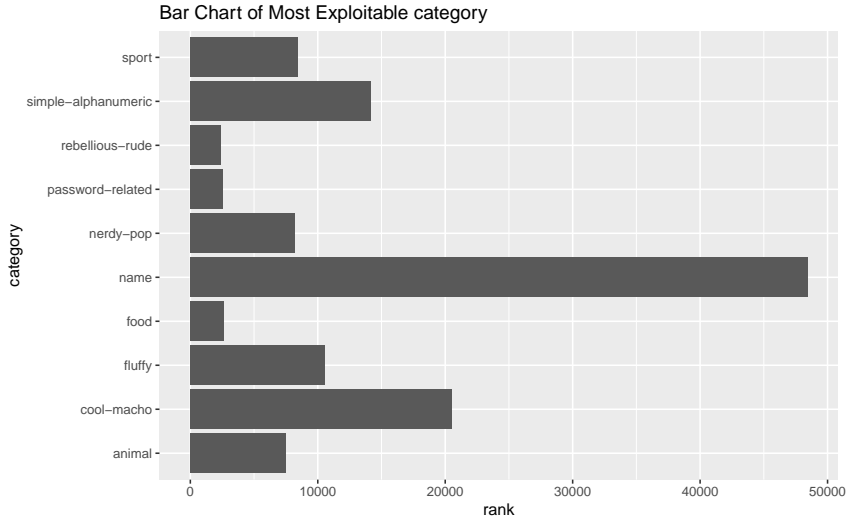
1. Most exploitable category are names.
2. Name will have the shortest time to be cracked.
3. There IS a relationship between rank and strength
4. There is a correlation between strength and value



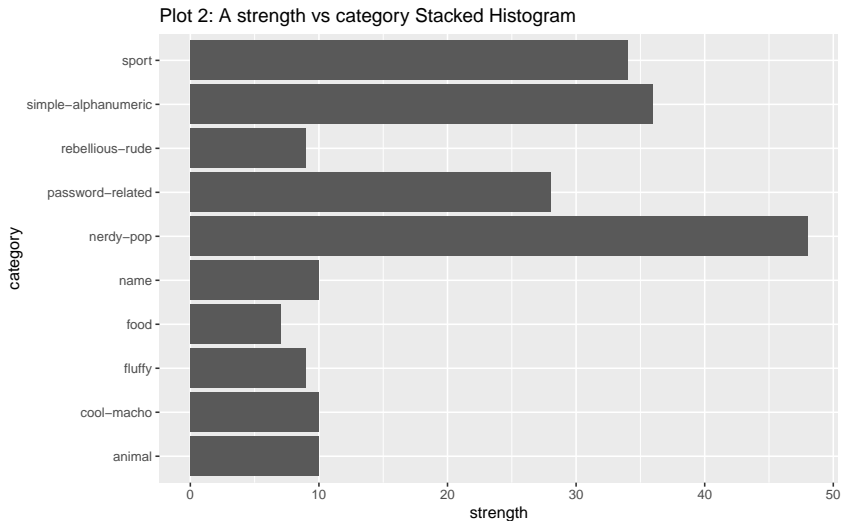
MACQUARIE
University
SYDNEY • AUSTRALIA

Datasets

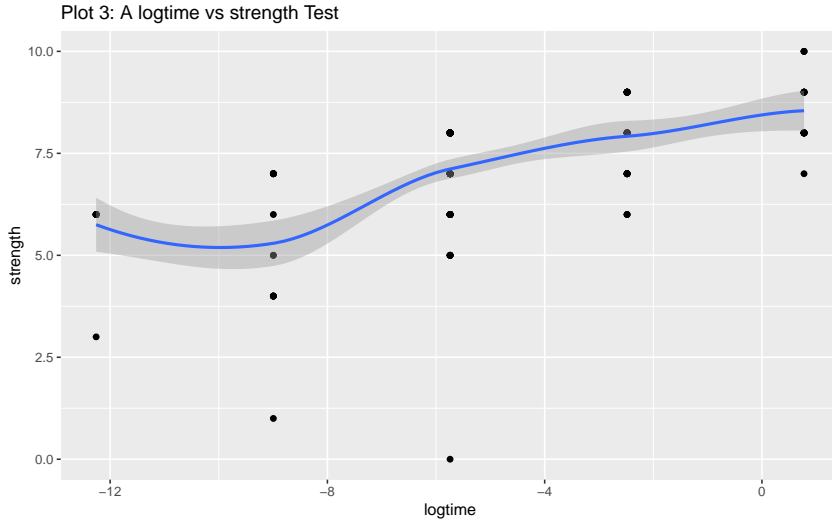
Most Exploitable Category.



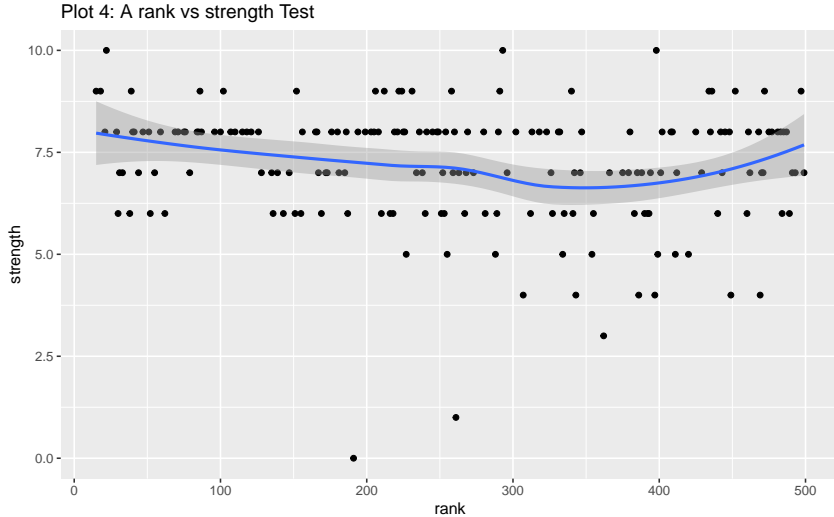
Category of the strongest and weakest passwords.



Category of the strength vs time to crack in s



Rank vs Strength ScatterPlot





MACQUARIE
University
SYDNEY • AUSTRALIA

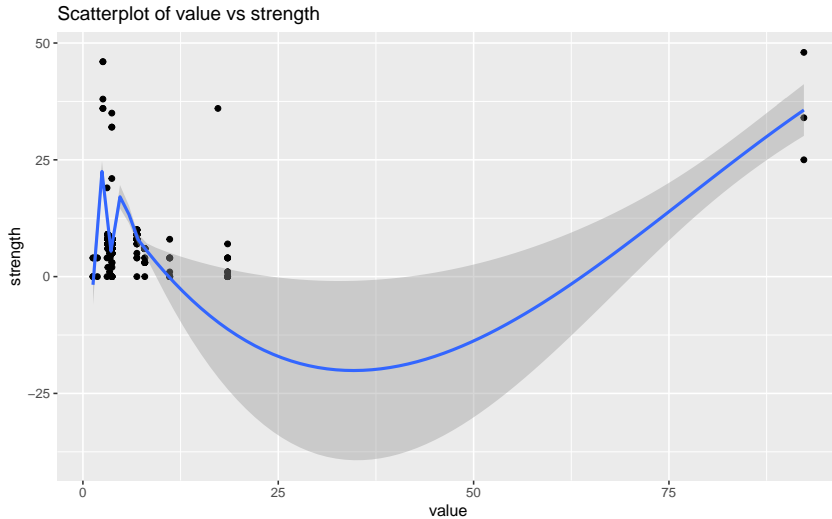
Correlation Test

```
correlation <- cor(x = passwordsFiltered$rank, y = passwordsFiltered$strength)
```

```
correlation
```

```
## [1] 0.0466191
```

Value vs Strength Scatter plot.



Conclusions

- ▶ Name was the most common password category.
- ▶ The time to crack some passwords took longer, regardless of what category.
- ▶ There IS NO linear regression between rank and strength.
- ▶ There IS a linear regression between strength and Time to crack.
- ▶ There is NO linear regression between value and strength.

Other Recommendations:

- ▶ Minimises chances of password being guessed with RNG password generation.
- ▶ Bitwarden changes the strength of the password.
- ▶ Through our observation, this will make more unique passwords that would not be categorized. As we realise, that passwords categorised as a name I.e. Fluttershy is more common than a RNG one.

- Grilo, Ferreira, and Almeida (2021) Suo, Zhu, and Owen (2005) Hughes (2020)
- Grilo, Miguel, João F Ferreira, and José Bacelar Almeida. 2021. "Towards Formal Verification of Password Generation Algorithms Used in Password Managers." *arXiv Preprint arXiv:2106.03626*.
- Hughes, Ellis. 2020. *tidytuesdayR: Access the Weekly 'TidyTuesday' Project Dataset*. <https://CRAN.R-project.org/package=tidytuesdayR>.
- Suo, Xiaoyuan, Ying Zhu, and G Scott Owen. 2005. "Graphical Passwords: A Survey." In *21st Annual Computer Security Applications Conference (ACSAC'05)*, 10–pp. IEEE.