# STAT545A Homework 1 repo of Rebecca Asiimwe

**Hi & Welcome to my STAT545 homework 1 repo.**

- A Quick Walk Through My Repo
- About Me
- Current Research
- Why STAT545
- The Awesomeness of R in a Nutshell
  - Useful Markdown Capabilities
    - Support for the LaTeX mathematical typesetting language
      - Supports intergration of code blocks
      - Markdown also supports intergration of youtube videos through links
- Fun Facts
  - Life Dares

- Follow Me

## A Quick Walk Through My Repo:

| File | Description |
|---|---|
| README.md | The readme.md file provides an overview of the ghist of my repo |
| hw01_gapminder.md | Data analysis and exploration of the gapminder dataset using R |
| plugins directory | |
| Maintains all auxillary files like images used for this repo | |

## About Me:

My name is *Rebecca Asiimwe*. I am a masters student in the MSc. Bioinformatics program at UBC. I come from an IT background that heavily revolved around working with computer networks,information systems, programming and working with databases; I have a Bachelor of Science degree in Information Technology and a Masters degree in Information Technology; however, I still seemed to be inclined to Biology!!!. My love for Biology drove me into Bioinformatics and I am loving it :smile:

## Current Research

I am currenly working on using data from whole genome seqeuncing to understand the underlying molecular and genomic underpinings behind Triple Negative Breast Cancer which is the most aggresive type of Breast Cancer.

My work also takes on a database driven approach of probing mutations and thereon conducting data analysis using python and R. I am currenlty based at the BC Cancer Agency and work in the Computational Biology

lab of Dr. Sohrab Shah

---

**Among the things we do in the lab involve understanding:**

1. *Cancer Evolution*
2. *Software development*
3. *Single cell genomics of cancer*
4. *Selection and drug response*
    - *Why do some patients respond to treatment and others dont?* ~~:wonder:~~ :thinking:
        - Well, we all have different genetic makeups but thanks to Next Generation Sequencing:
            * we are begining to understand complex diseases and advances into finding drugable targets for non responsive patients are underway with some major breakthroughs reported in some cancer types like HER-2 dependant breast cancers.
        - The tumor-micro environemts (TME) differ and also have a large bearing on tumor evolution
            1. Can we use machine learning to understand the TME?
            2. Can the TME throw more light on disease progression?
5. *Mutational signatures in DNA repair deficient cancers, signatures like:*
    1. APOBEC
    2. HRD
    3. POLE
    4. FBI

---

## Why STAT545

Over the past year, I have grown a strong love and desire for data wrangling, exploration, and analysis both in R and Python :ok_hand:. I took STAT545 as a way of introducing me to key concepts required for me to succeed as a data analyst and scientist. The data I work with is high dimensional biological data and getting exposed to the concepts required to work with such datasets will go a long way in helping me analyse and interpret this data.

## The Awesomeness of R in a Nutshell

| | | | |
|---|---|---|---|
| Wide usage in academia; enourmous availability of resources; active coding community; freely available as an open source tool | Awesome at data wrangling! | Great with data visualization | Provides ample tools for machine learning |

## Useful R Markdown Capabilities

### Support for the LaTeX mathematical typesetting language

Produces nice-looking documents with R input and output neatly formatted. Examples hereunder:

### Supports intergration of code blocks

Example Code:

We can also cite code from `R` and other programming languages like `Java, C++ and Python`:- Below is a ***piece of code cited from python***

```
import pandas as pd

gapminder=pd.read_pickle('/Users/rasiimwe/gapminder.pkl')
gapminder.head(4)
gapminder.head.index=gapminder[gapminder.columns[0]]
# ... more data wrangling lines of code using pandas
gap_transpose=gap_transpose.T

year=2007
plt.hst(gap_transpose.ix[year].dropna().values, bins=10)
```

```
plt.xlabel("gdp per capita")
Plt.ylabel(count)
```

**Markdown also supports intergration of youtube videos through links**

Check out these interesting videos:

**R's posibilities into data visualization** **Basketball - Epic fake plays – simply awesome**

**A programmer's life!** **Hitler uses git!!!!!!!!** :laughing:

---

Before I bolt out of here :paw_prints: :paw_prints: :paw_prints: `some fun facts` :point_down:

## Fun facts

:dancer: :dancer: :dancer: :dancer:

**I love:** :guitar: :musical_keyboard: volleyball :volleyball: :headphones: :microphone: :musical_note: :horse_racing: :airplane: – cooking, baking, singing, watching movies, hanging out with friends . . .

I am not a fun of sea food but I want to try :octopus: :wink:

My friends call me "smiling machine" :blush: :blush:

**Life dares**

- [ ] Mountain climbing
- [ ] Horse riding
- [ ] Eating strange or weird foods like scorpion :stuck_out_tongue_winking_eye:
- [x] Try new colors
- [x] Choose to be joyful regardless :dancer: :dancer:

---

***You can follow me on:***