

# Contents

<b>Home</b>	<b>3</b>
0.1 Latest Projects . . . . .	3
<b>About Me</b>	<b>5</b>
<b>1 Data Exploration with Shiny</b>	<b>7</b>
<b>2 Useful Links</b>	<b>9</b>
<b>3 Applications</b>	<b>11</b>
3.1 Example one . . . . .	11
3.2 Example two . . . . .	11
<b>4 Final Words</b>	<b>13</b>



# Home

Welcome to my blog! This is where I share excerpts of my portfolio as well as professional resources I am currently using. I will also share some blog posts about experiences from my working life.

## 0.1 Latest Projects

- [Building a Data Exploration Web App with Shiny][data-exploration-using-shiny]



# About Me

Jillian Augustine, PhD.

I am a data scientist and question answerer. I received my PhD in Molecular Biology from the University of Vienna (Austria) and my BSc from the University of Leeds (UK) during which I also studied at McGill University (Montreal, Canada).

My main interest (and the reason I became a data scientist) is using data to answer questions regardless of the industry. I have experience working in both telecommunications and the Paper & Packaging industries. In particular I am passionate about data understanding and communication both to stakeholders and within data teams. My professional approach to data science is to use the tool that gets the job done given any constraints from team members and stakeholders. Here is a list of my current go-to tools.

Data Manipulation:

- python
- Apache Spark
- R
- any combination of the above
- Microsoft Excel (for team members accustomed to pivot tables)

(Real-Time) Data Collection/Extraction:

- SQL (Hive)
- Apache Kafka

Data Storage:

- parquet files whenever possible
- Hadoop Distributed File System (HDFS)
- Linux computer clusters

Data Visualisation:

- ggplot2

Machine Learning:

- caret
- (scikit-learn)

Documentation/Project Work:

- R Markdown
- Confluence
- Jira
- Git

Another interest of mine is increasing the inclusivity of working groups through open exchanges and active diversification. I try to make my presentations as accessible as possible and welcome feedback as to how

I can improve this further. I welcome opportunities to speak about my work at conferences and meetups. Please get in touch through Twitter or LinkedIn if you would like to know more.

# Chapter 1

## Data Exploration with Shiny

Why?

How?

My ideas - email notification

Requirements from users?

Requirements from the business - logging of users requests and viewing - automatic deletion of files - limited access to data

Prototypes?





## Chapter 2

# Useful Links

I am often asked about which resources I use to improve and maintain my data science skills. Here is a list of websites I often use and would recommend.



## Chapter 3

# Applications

Some *significant* applications are demonstrated in this chapter.

### 3.1 Example one

### 3.2 Example two



## Chapter 4

# Final Words

We have finished a nice book.