

Data Science for the Juvenile Court

Martin Gleason, MS

12/5/2017

Ubiquity, Abstraction, Absurdity



Figure 1:

Human Driven Process

SHEET NO. 1

ACCOUNT NO. 20

TERMS. Pay Barten 5⁰⁰ v x 1/2 NAME Jo Pains Yellow & Wife

RATING. 8⁰⁰ - 1⁵⁰ ADDRESS

CREDIT LIMIT.

DATE 1815	ITEMS	Folio	✓	DEBITS	DATE	ITEMS	Folio	✓	CREDITS
Nov 20	Barten Contract			40 -	Nov 13	Salary	11		9.50
20	Strong Curial			11.50	13				35 -
13		11		4.50					11.50
				56 -					56 -
13	Btw Bal.			35 -	20	"	12		9.50
13	WAB "			11.50					30
20		12		4.50					11.50
20	Btw "			30 -	27		13		9.50
20	WAB "			11.50					25 -
27		13		4.50					11.50
27	Btw "			25 -	Dec 4		14		9.50
27	WAB "			11.50					20 -
Dec 4		14		4.50					11.50
11	Btw "			20 -	11		15		9.50
11	WAB "			11.50					15
11		15		7.75					11.75
11	Btw "			15 -	18		16		9.50
11	WAB "			11.75					10 -
									10.25
18	Btw "			10 -	20		17		9.50
18	WAB "			10.25					51
									57.5
				20.25					20.25

Figure 2: Spreadsheets are digital ledgers

Complex But Concrete

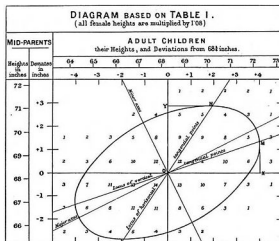


Figure 3: Galton's Work on Regression to the mean. He was also a racist piece of garbage

Cloud Computing



Figure 4: Complex AND Abstract

Data science, what is it?

- ▶ Data Science is the way we study the capture, storage, and transformation of data into information.
- ▶ It's that complex stuff Marty talks about all the damn time.
–Toastmasters, 2017

Data Scientists



Hadley Wickham,
PhD

Data Science is



Figure 6: Using data to solve problems

What is Social Good?



Figure 7: If we work at 1100 S Hamilton, we work for the social good

What is Data Science for the Social Good?

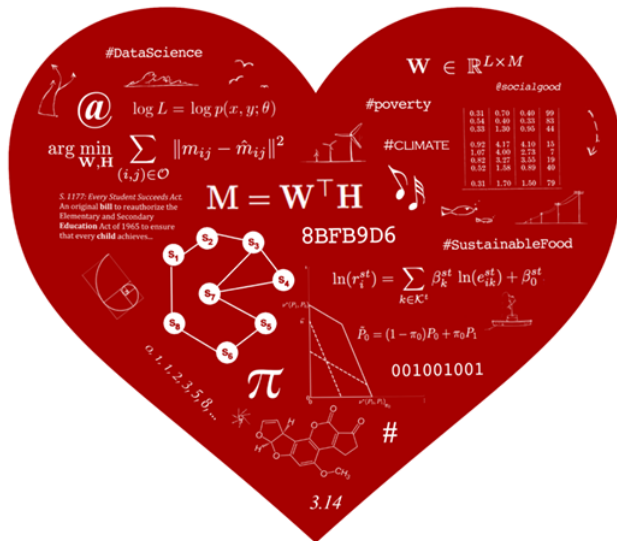
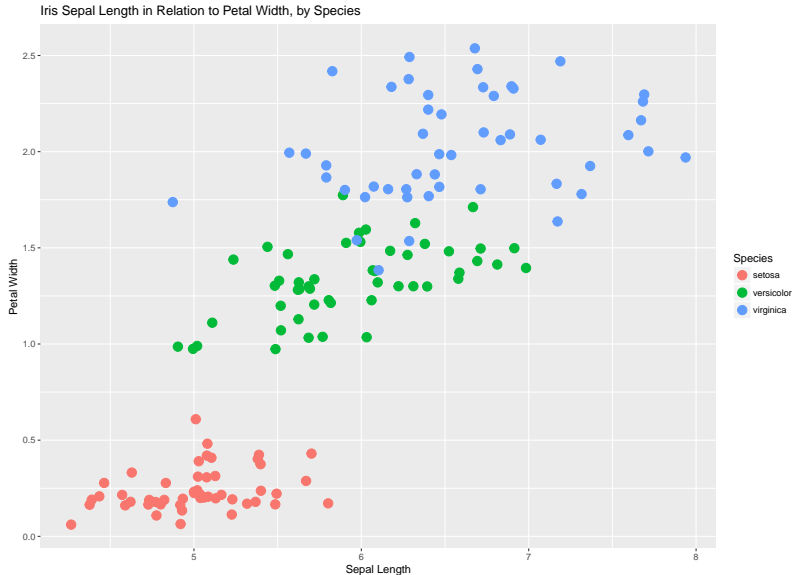


Figure 8:

Data Science for the Social Good

- ▶ Using data science principles to make the world a better place
- ▶ Identifying problems
- ▶ New Solutions
- ▶ Reproducing it to scale
- ▶ What would Data Science for the Juvenile Court Look Like?

Concrete Steps: What is Visualizing our Information?



Sepal Plot

- ▶ Shows us at a glance what the relationship is between flower parts across three species of flower
- ▶ What does our data show us?

Our Data, Visualized

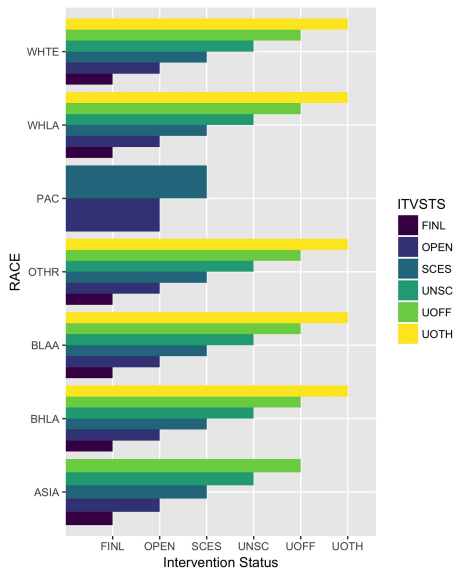


Figure 9: Based on Diversion Data from 2012-2016

Methods

- ▶ Tidy Data
- ▶ Statistical Analysis
- ▶ Understanding the problem

The screenshot displays the RStudio interface. The main editor window contains R code for data science, including comments about the Cour 101/DataScience project and the use of the tidyverse. The console window shows the R version (3.4.2) and the R Foundation for Statistical Computing license. The sidebar on the right displays the 'The tidy tools manifesto' document, which outlines the principles of the tidyverse.

```

19 </div>
20 (Personal)/codingProjects/R/DataScience for the Court 101/DataOutlook.jpg)
21 ## Human Driven Process
22 <div align = "center">
23 |Sproadsheetes are digital ledgers|C:/Users/marty/Dropbox (Personal)/codingProjects/R/DataScience for the Court 101/Ledger.jpg)
24 </div>
25
26 ## Complex But Concrete
27 <div align = "center">
28 |Galton's Work on Regression to the mean. He was also a racist piece of garbage|C:/Users/marty/Dropbox
(Personal)/codingProjects/R/DataScience for the Court 101/Galton's correlation diagram.1875.jpg)
29 </div>
30 ## Cloud Computing
31 <div align = "center">
32 |Complex AND Abstract|C:/Users/marty/Dropbox (Personal)/codingProjects/R/DataScience for the Court 101/cloudcomputing.png)
33 </div>
34
35 ## Data science, what is it?
36 <div align = "center">
37 |C:/Users/marty/Dropbox (Personal)/codingProjects/R/DataScience for the Court 101/datascience.jpg)
38 </div>
39
40 ## Data science, what is it?
41 * Data Science is the way we study the capture, storage, and transformation of data into information.
42
43 * It's that complex stuff Marty talks about all the damn time. --Toastmasters, 2017
44
45 ## Data Scientists
46 |Data science is an exciting discipline that allows you to turn raw data into understanding, insight, and
knowledge.|C:/Users/marty/Dropbox (Personal)/codingProjects/R/DataScience for the Court 101/hadley.jpg) --Hadley Wickham, PhD
24.8 Human Driven Process

```

Console Terminal R Markdown

```

~/Dropbox (Personal)/codingProjects/R/DataScience for the Court 101 /
R version 3.4.2 (2017-09-28) -- "Short Summer"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'licence()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/Dropbox (Personal)/codingProjects/R/DataScience for the Court 101 /RData]
>

```

Environment History Connections

```

library(ggplot2)
geom_label(aes(x = "Sepal Length", y = "Petal Width", title = "Iris Sepal...
ggplot(Iris, aes(x = Sepal.Length, y = Petal.Width, col = Species)) + geo...
labs(x = "Sepal Length", y = "Petal Width", title = "Iris Sepal length in...
heart <- image_read("Users/marty/Dropbox (Personal)/codingProjects/R/Dat...
image_write(heart)
image_display(heart)
heart
knitr::opts_chunk$set(echo = FALSE)
library("tidyverse")
library("magick")
knitr::knitr()
heart <- image_read("Users/marty/Dropbox (Personal)/codingProjects/R/Dat...
knitr::opts_chunk$set(echo = FALSE)
library("tidyverse")
library("magick")
library("knitr")
heart <- image_read("Users/marty/Dropbox (Personal)/codingProjects/R/Dat...

```

Files Plots Packages Help Viewer

The tidy tools manifesto - Find in Topics

The tidy tools manifesto

Hadley Wickham

2017-11-13

This document lays out the consistent principles that unify the packages in the tidyverse. The goal of these principles is to provide a uniform interface so that tidyverse packages work together naturally, and once you've mastered one, you have a head start on mastering the others.

This is my first attempt at writing down these principles. That means that this manifesto is both aspirational and likely to change heavily in the future. Currently no packages precisely meet the design goals, and while the underlying ideas are stable, I expect their expression in prose will change substantially as I struggle to make explicit my process and thinking.

There are many other excellent packages that are not part of the tidyverse, because they are designed with a different set of underlying principles. This doesn't make them better or worse, just different. In other words, the complement to the tidyverse is not the messyverse, but many other universes of interrelated packages.

There are four basic principles to a tidy API:

1. Reuse existing data structures.
2. Compose simple functions with the pipe.
3. Embrace functional programming.
4. Design for humans.

Figure 10: The code behind this presentation

Data Science for the Juvenile Court

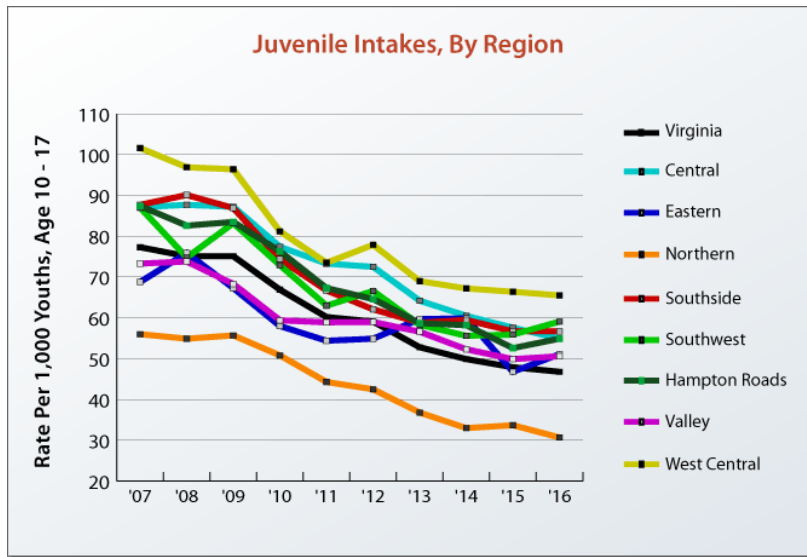


Figure 11: End Game: Explaining this!