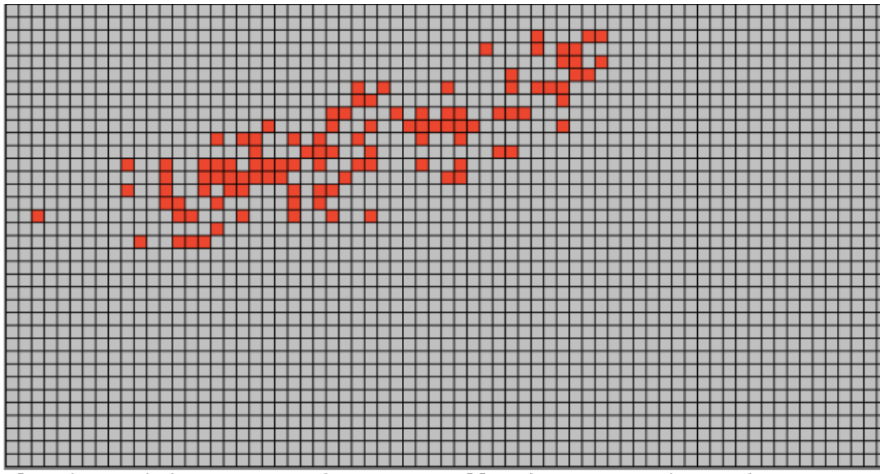
Below are some of the coolest project 7 submissions. We have broken them up into different categories based on the type of analysis performed by the creative methods. Many thanks to all of you for the effort and creativity you have put into this project.

GRAPHS OF CORRELATION COEFFICIENTS

Corrine Farley

Analyzed the college data set by calculating correlation coefficients between various metrics and creating visualizations of the correlations using grids.



```
Plotting tuitionOutState (thousand dollars) vs. actScience (score out
of 36)
Correlation Coefficient r = 0.8179428345830893
```

Daniel Song

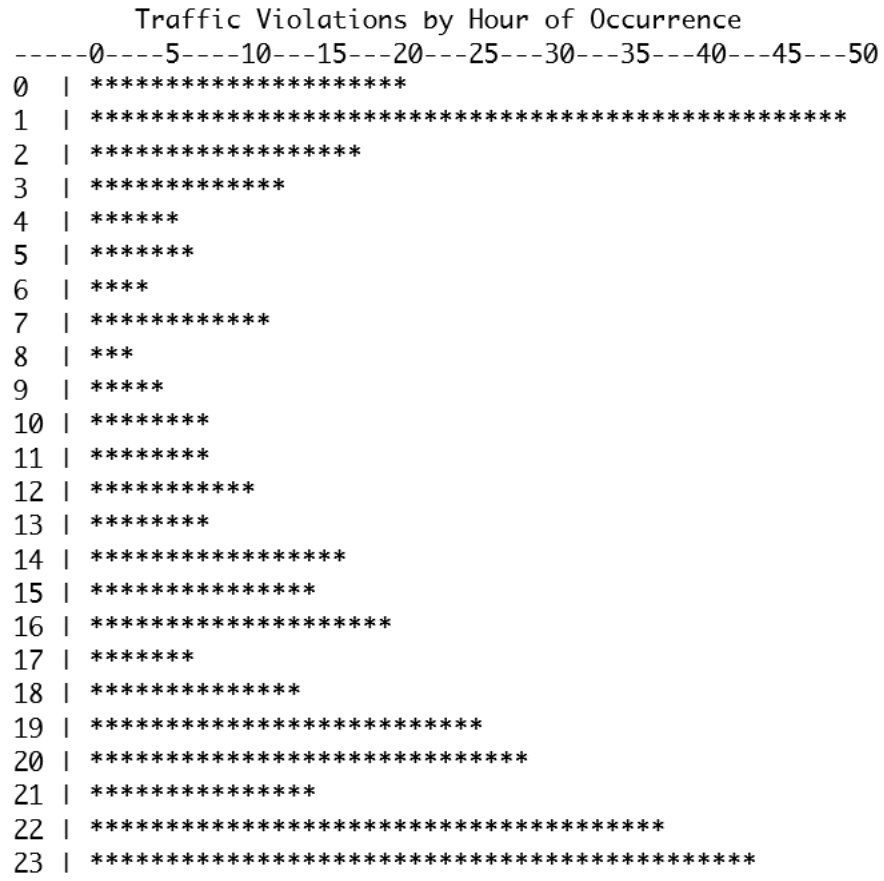
The same information as above, but with the MLB data set and ASCII graphics rather than a grid.

```
BATTINGAVERAGE units in 0.01 vs. SINGLES, units in 10.0
Correlation between battingAverage & singles: 67.21%
 37|
 36|         *                         *
 35|      *
 34|
 33|                                   *      *   *
 32|                  *          *   *   *   *   *   *
 31|              *          *   *   *   *   *   *      *          *
 30|          *   *      *   *   *   *   *   *   *   *   *
 29|   *      *   *      *   *   *   *   *   *   *   *
 28|      *   *      *   *   *   *   *   *   *   *   *   *
 27|          *   *   *   *   *   *   *   *   *   *   *
 26|   *   *   *   *   *   *   *   *   *   *
 25|   *   *   *   *   *   *   *   *   *   *
 24|       *   *   *   *   *   *       *   *
 23|   *   *   *   *   *   *   *   *
 22|   *   *   *   *   *   *
 21|       *   *   *   *   *
 20|   *   *   *
 19|   *   *       *
 18|   *
 17|   *       *
 16|   *   *
 15|   *
 14|
 13|   *
 12|
 11|
 10|
  9|
  8|
  7|
  6|
  5|
  4|
  3|
  2|
  1|
  0|
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19
```

ASCII HISTOGRAMS OF TRAFFIC DATA

Matthew Bizet

Simple, but very neatly formatted bar graph of traffic stops per hour of the day.



The next two projects show more information, but are perhaps harder to interpret, showing the balance between clarity and information density.

Samina Hussain

Visualization of traffic stops in the form of a histogram, but neatly formatted and including information about both car color and gender in one plot.

```
Histogram
BEIGE: +------- (11% F, 88% M)
BLACK: ++++++++++++++++++++++++++++++++-------- (44% F, 55% M)
BLUE: ++++++++++++---------- (29% F, 70% M)
BLUE DARK: ++---------- (14% F, 85% M)
GOLD: +++-- (37% F, 62% M)
GRAY: ++++++++++++++++++++++---------- (38% F, 61% M)
GREEN: ----- (0% F, 100% M)
GREEN DK: -- (0% F, 100% M)
MAROON: ++-- (33% F, 66% M)
PURPLE: - (0% F, 100% M)
RED: +++---------- (15% F, 84% M)
SILVER: ++++++++++++++++++++++++++++++++------------------------- (36% F, 63% M)
TAN: +- (33% F, 66% M)
WHITE: ++++++++++++++++++++++++++----------- (39% F, 60% M)
YELLOW: - (0% F, 100% M)

- = male, + = female, # = equal
```

Ruth Shatkay

Visualizing the same information (number of stops by car color and gender) and using an ASCII method, but with a completely different visualization. Colors are represented by different characters, and males and females are split onto different lines instead of making up one whole bar.

```
KEY: beige - '*', black - ':', blue - '^', blueDark - '.', gold - '!', gray -
'#', green - '+', greenDk - '-', maroon - 'm', na - 'n', orange - 'o', purple
- 'p', red - 'r', silver - 's', tan - 't', white - 'w', yellow - 'y'
*
********
::::::::::::::::::::::::::::::::
::::::::::::::::::::::::::::::::::::::::
^^^^^^^^^^^^
^^^^^^^^^^^^^^^^^^^^^^^^^^^^

..
............
!!!
!!!!!
########################
########################################

++++++

--
mm
mmmm

n

o

p
rrr
rrrrrrrrrrrrrrrr
t
tt
##################################
############################################################
wwwwwwwwwwwwwwwwwwwwwwwww
wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww

y
```

USEFUL TOOLS FOR AN END-USER

Yannick Zanfack

If a user specifies a state, the code summarizes the options for colleges in that state. If the user specifies an ACT score, the code returns a list of "good fit" colleges with similar ACT averages.

The code also identifies which states have a larger difference between in-state and out-of-state tuition. Very useful for a prospective student.

The first method is called **collegeOptionsByState( )** and has the following output :

```
You have the option between 3 Colleges in MA
Boston College ==> Out of state tuition : $45622 ; In state tuition : $45622
Harvard University ==> Out of state tuition : $42292 ; In state tuition : $42292
University of Massachusetts-Lowell ==> Out of state tuition : $26146 ; In state tuition : $12097
The average out of state tuition in MA is : $38020.0
The average in state tuition in MA is : $33337.0
```

The second method ***compareActScore ( )*** has the following output :

```
Your average ACT score is : 19.0
You meet the requirements for the following Colleges :
University of Arkansas at Pine Bluff,  Average ACT 17.666666666666668
Indiana State University,  Average ACT 18.666666666666668
Morgan State University,  Average ACT 17.666666666666668
University of Maryland Eastern Shore,  Average ACT 17.0
Texas A & M International University,  Average ACT 18.333333333333332
Texas A & M University-Kingsville,  Average ACT 18.666666666666668
```

Tumasang Che

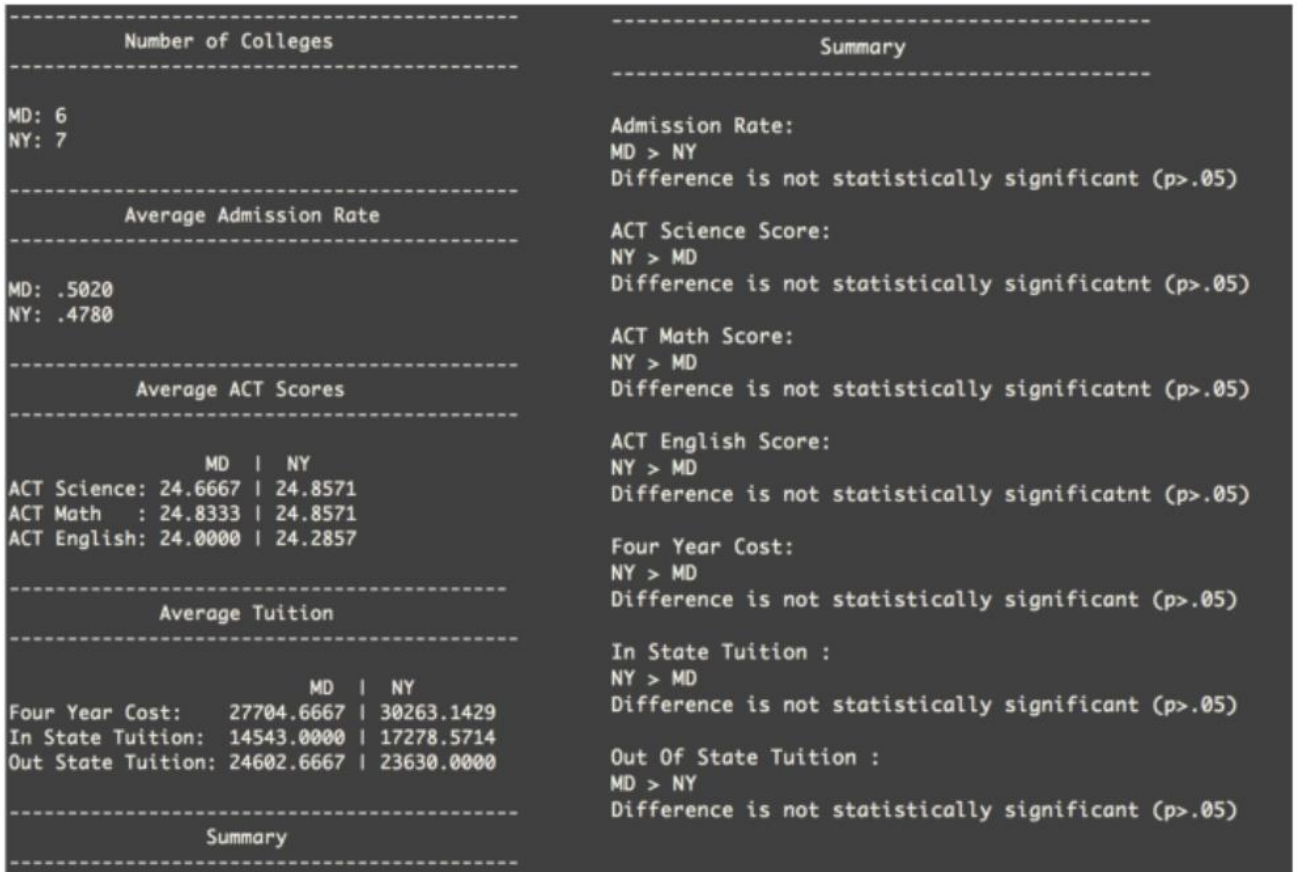
Calculates the slugging average of each individual player, followed by a team rating which averages the members of a given team. In order to inform team managers which players should be traded due to poor performance, a “transfer market” shows the information of those players who performed significantly worse than the team.

Players in transfer Market: 43

id	playerName	salary	position	atBats	hits	singles	doubles	triples	homeRuns	battingAverage
8	Luis Castillo	6250000	4	298	73	62	7	1	3	0.245
9	Marlon Anderson	1050000	11	138	29	22	6	0	1	0.21
11	Rob Mackowiak	1500000	11	53	7	5	1	0	1	0.132
35	Rob Bowen	410000	2	91	16	9	5	1	1	0.176
38	Ronny Paulino	423500	2	118	25	18	5	0	2	0.212
44	Morgan Ensberg	1750000	5	74	15	14	0	0	1	0.203
46	Andruw Jones	14726910	8	209	33	21	8	1	3	0.158
55	Corky Miller	400000	2	60	5	4	0	0	1	0.083
73	Carlos Ruiz	425000	2	320	70	52	14	0	4	0.219
78	So Taguchi	900000	7	91	20	14	5	1	0	0.22
79	Craig Counsell	2800000	5	248	56	40	14	1	1	0.226
89	Josh Bard	2237500	2	178	36	26	9	0	1	0.202
90	Ben Broussard	3850000	3	82	13	10	0	0	3	0.159
92	John McDonald	1900000	6	186	39	30	8	0	1	0.21
99	Chris Burke	955000	11	165	32	24	5	1	2	0.194
106	Eric Bruntlett	600000	6	212	46	34	9	1	2	0.217
110	Donnie Murphy	400000	5	103	19	13	3	0	3	0.184
127	Wily Mo Pena	2000000	7	195	40	32	6	0	2	0.205
159	Paul Lo Duca	5000000	2	173	42	33	9	0	0	0.243
160	Michael Barrett	3500000	2	94	19	14	3	0	2	0.202
173	Willie Bloomquist	1000000	8	165	46	45	1	0	0	0.279
186	Robby Hammock	409500	2	42	8	7	1	0	0	0.19
191	Guillermo Quiroz	400000	2	134	25	18	5	0	2	0.187
195	Tadahito Iguchi	3850000	4	310	72	54	15	1	2	0.232
196	Tony Pena	405500	6	225	38	32	4	1	1	0.169
204	Augie Ojeda	550000	4	231	56	45	9	2	0	0.242
216	Shannon Stewart	1500000	7	175	42	35	4	2	1	0.24
218	Omar Vizquel	5000000	6	266	59	48	10	1	0	0.222
238	Humberto Quintero	405000	2	168	38	30	6	0	2	0.226
239	Willy Taveras	1975000	8	479	120	102	15	2	1	0.251
243	Brad Ausmus	2000000	2	216	47	36	8	0	3	0.218

Jee Kang

Gives a thorough comparison on colleges in two states. Very nicely formatted output, useful for a student looking for potential schools.



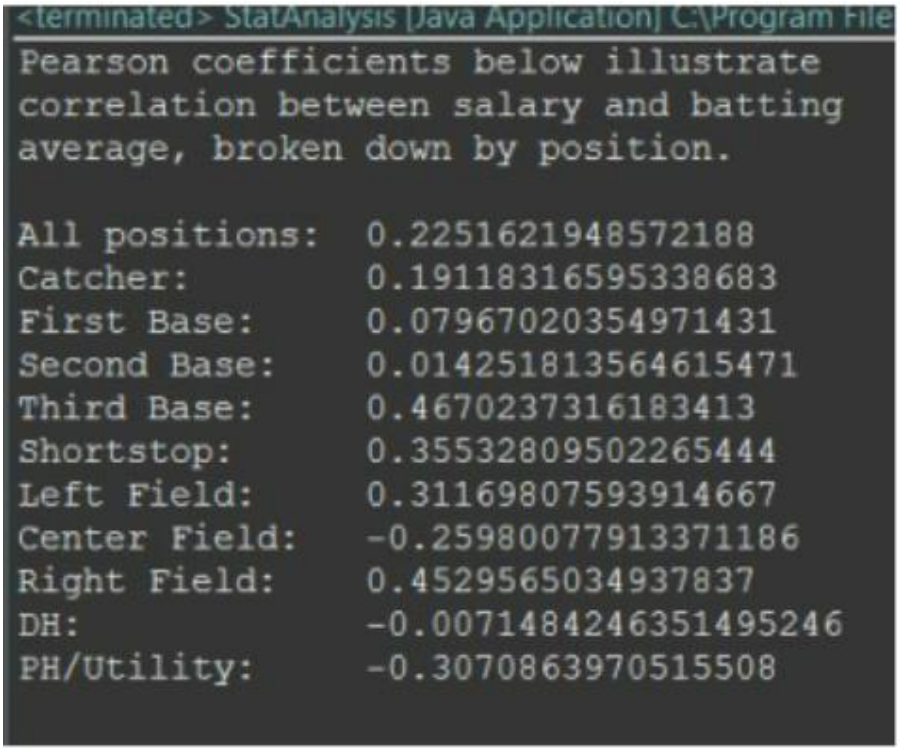
```
-----------------------------------------
         Number of Colleges
-----------------------------------------

MD: 6
NY: 7

-----------------------------------------
         Average Admission Rate
-----------------------------------------

MD: .5020
NY: .4780

-----------------------------------------
          Average ACT Scores
-----------------------------------------

                MD   |  NY
ACT Science: 24.6667 | 24.8571
ACT Math   : 24.8333 | 24.8571
ACT English: 24.0000 | 24.2857

-----------------------------------------
            Average Tuition
-----------------------------------------

                        MD   |  NY
Four Year Cost:    27704.6667 | 30263.1429
In State Tuition:  14543.0000 | 17278.5714
Out State Tuition: 24602.6667 | 23630.0000

-----------------------------------------
               Summary
-----------------------------------------
```

```
-----------------------------------------
                Summary
-----------------------------------------

Admission Rate:
MD > NY
Difference is not statistically significant (p>.05)

ACT Science Score:
NY > MD
Difference is not statistically significatnt (p>.05)

ACT Math Score:
NY > MD
Difference is not statistically significatnt (p>.05)

ACT English Score:
NY > MD
Difference is not statistically significatnt (p>.05)

Four Year Cost:
NY > MD
Difference is not statistically significant (p>.05)

In State Tuition :
NY > MD
Difference is not statistically significant (p>.05)

Out Of State Tuition :
MD > NY
Difference is not statistically significant (p>.05)
```

Brian Daisey

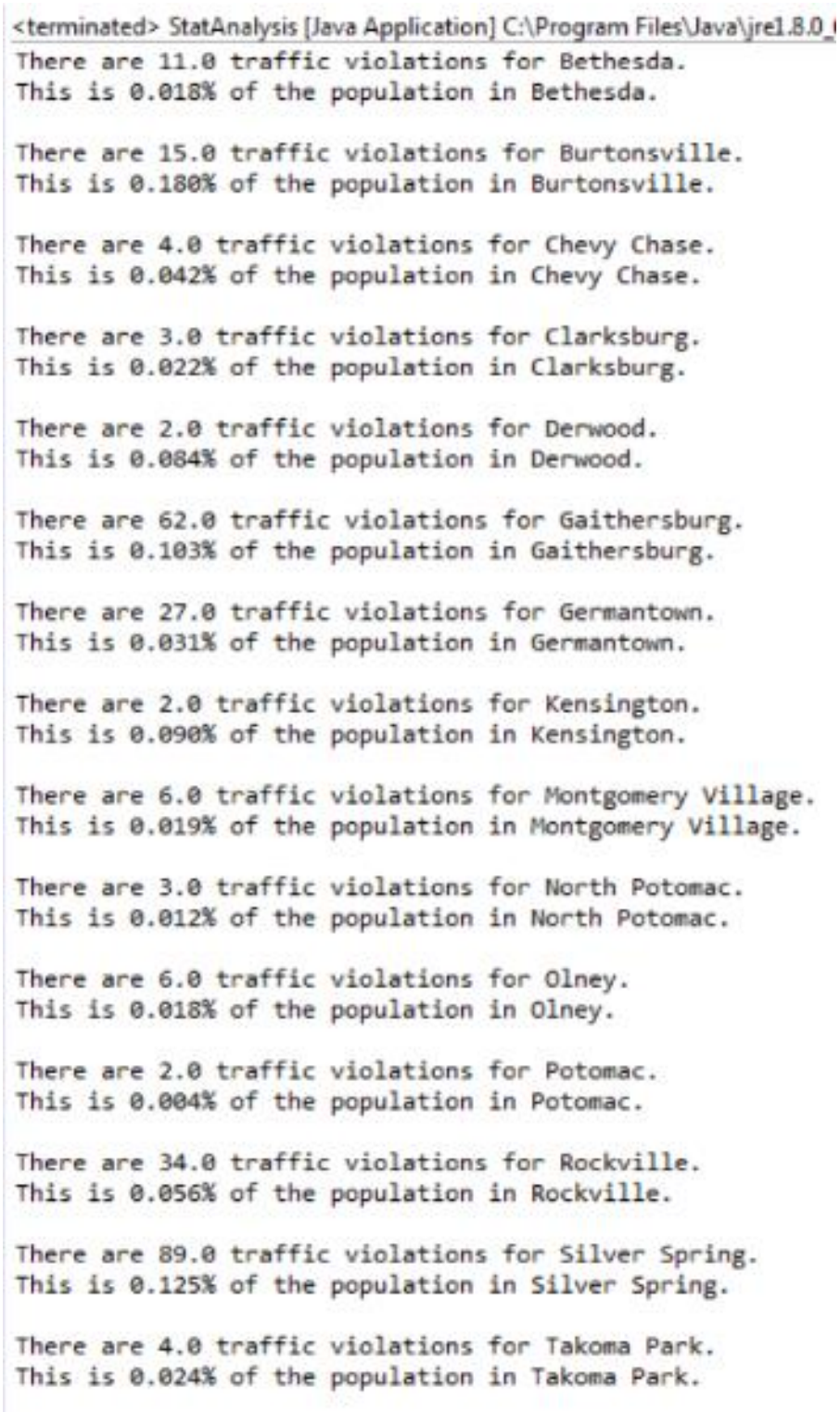
Used Pearson Correlation Coefficient to determine whether there is a significant correlation between salary and batting average, with data facetted by position. Developed hypotheses and analyzed each finding thoroughly. Interesting findings: players in similar positions have very different correlation coefficients, and even designated hitters do not seem to get paid more for having higher batting averages.

Exceptionally well-written as a research-style report.



Michelle Decaire

Great use of external data to supplement the data provided in the project. Pulled data from the 2010 US census to determine the population of each area of the county. Cross-referenced number of traffic stops per area with total population to find the rate at which people were pulled over in each area.



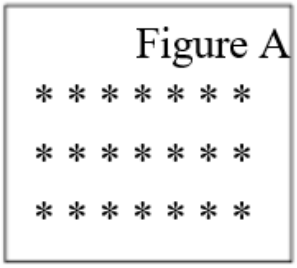
David Fisher

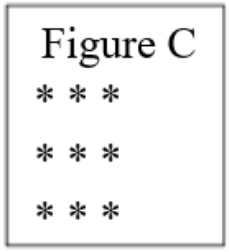
Also a great example of the use of external data. Analysis of baseball players by a variety of metrics (accurate to actual MLB data) such as “weighted on-base average” for hitters and “fielding independent pitching” for pitchers, player position, and player league (National or American). Created an optimal all-star team for each league using these metrics, then compared it to the actual all-star roster to find that 53% of the players he selected were on the actual roster. Below, green indicates exact roster position correct and yellow indicates presence on roster correct.

	AllStar.java		Real 2008 All Star Roster	
	National League	American League	National League	American League
C	McCann. B	Mauer. J	Geovany Soto	Joe Mauer
1B	Pujols. A	Teixeira. M	Lance Berkman	Kevin Youkilis
2B	Utley. C	Kinsler. I	Chase Utley	Dustin Pedroia
SS	Ramirez. H	Jeter. D	Hanley Ramirez	Derek Jeter
3B	Jones. C	Rodriguez. A	Chipper Jones	Alex Rodriguez
OF	Holliday. M	Ramirez. M	Ryan Braun	Josh Hamilton
OF	Ludwick. R	Quentin. C	Kosuke Fukudome	Manny Ramirez
OF	Dunn. A	Markakis. N	Matt Holliday	Ichiro Suzuki
OF/DH	n/a	n/a	Alfonso Soriano(injured)	David Ortiz(injured)
P	Lincecum. T	Lee. C	Aaron Cook	Justin Duchscherer
P	Sabathia. C	Halladay. R	Ryan Dempster	Roy Halladay
P	Haren. D	Beckett. J	Dan Haren	Scott Kazmir
P	Lowe. D	Santana. E	Brad Lidge	Cliff Lee
P	Webb. B	Mussina. M	Tim Lincecum	Joe Nathan
P	Billingsley. C	Danks. J	Carlos Marmol	Jonathan Papelbon
P	Lidge. B	Rivera. M	Ben Sheets	Mariano Rivera
P	Saito. T	Papelbon. J	Edinson Volquez	Francisco Rodriguez
P	Rhodes. A	Ramirez. R	Billy Wagner	Ervin Santana
P	Wood. K	Thornton. M	Brandon Webb	Joe Saunders
P	Fuentes. B	Nathan. J	Brian Wilson	George Sherrill
P	Sheets. B	Burnett A.	Kerry Wood	Joakim Soria
P	Dempster. R	n/a	Carlos Zambrano	n/a
Reserve				
C	Soto. G	Suzuki. K	Russell Martin	Dioner Navarro
C	Martin. R	Pierzynski. A	Brian McCann	Jason Varitek
1B	Berkman. L	Youkilis. K	Adrian Gonzalez	Justin Morneau
1B	Fielder. P	Pena. C	Albert Pujols	Ian Kinsler
2B	Uggla. D	Roberts. B	Dan Uggla	n/a
2B	deRosa. M	Pedroia. D		n/a
SS	Wright. D	Ramirez. A	Cristian Guzman	Michael Young
SS	Reyes. J	n/a	Miguel Tejada	n/a
SS	Hardy. J	n/a		n/a
3B	Drew. S	Longoria. E	David Wright	Joe Crede
3B	Ramirez. A	Gordon. A	n/a	Carlos Guillen
3B	n/a	Scutaro. M	n/a	Evan Longoria
OF	Burrell. P	Sizemore. G	Corey Hart	J.D. Drew
OF	Beltran. C	Bay. J	Ryan Ludwick	Carlos Quentin
OF	Hawpe. B	Hamilton. J	Nate McLouth	Grady Sizemore
DH	n/a	n/a	n/a	Milton Bradley

Andrew Cachiaras

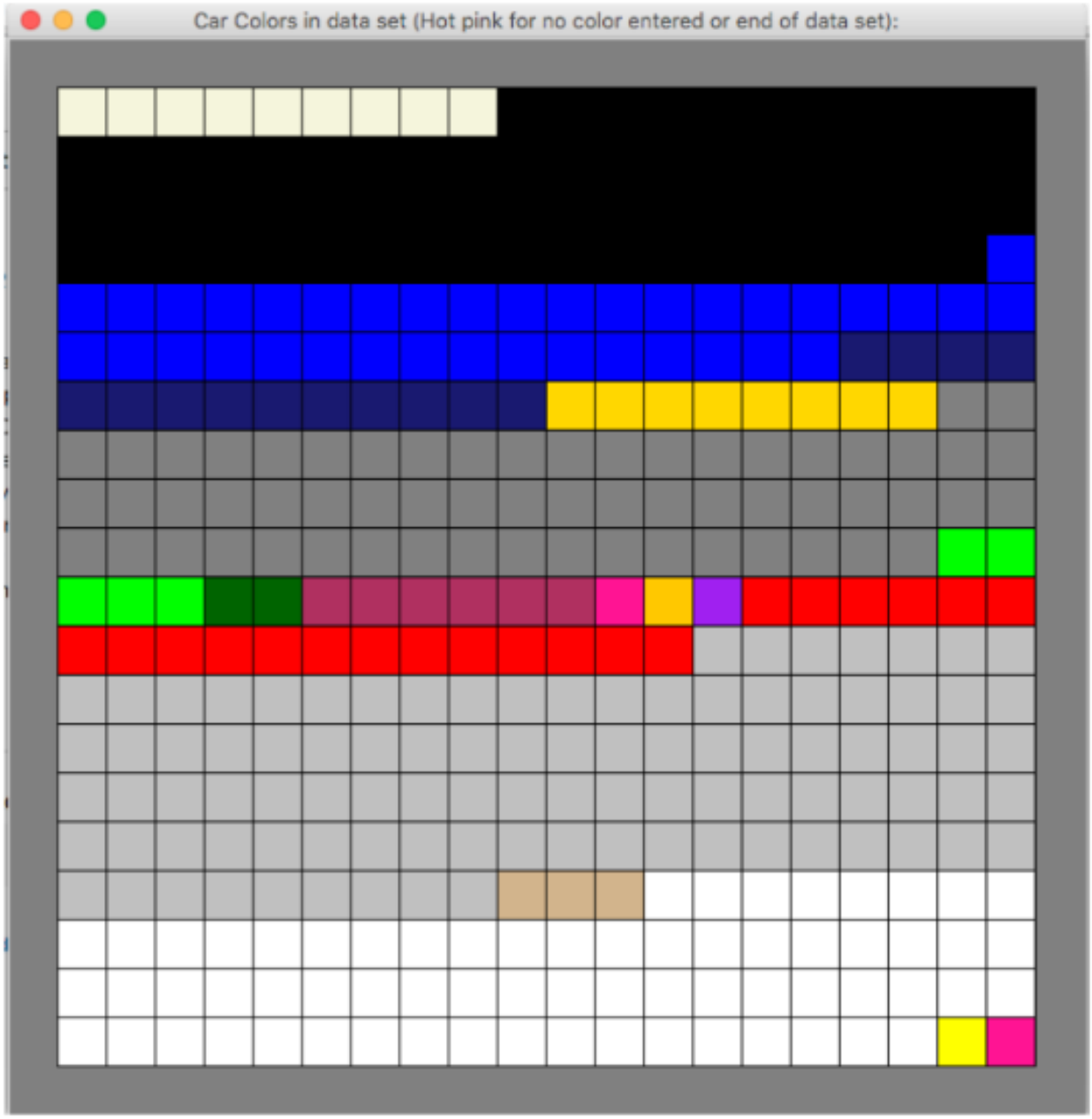
An example of an interesting/creative visualization approach. Created a visual representation of the correlation between salary and hitting quality in order to show which players are overpaid/bad investment (Figure A, wide and short rectangle), underpaid/good investment (Figure B, tall and thin rectangle), or appropriately paid (Figure C, square). A non-technical user (e.g., baseball team manager) can very quickly understand the data from this visualization.





Kevin Doherty

Counted the number of traffic stops for each car color, translated the colors to RGB equivalents, and created a grid visualization of the traffic stops in the data set. There is no effort required to interpret the information, which makes the data immediately understandable without having to look through a long CSV file.



Colors in order: Beige, Black, Blue, Blue Dark, Gold, Gray, Green, Green DK, Maroon, “N/A” - read as default value set for null colors. Orange, Purple, Red, Silver, Tan, White, Yellow.

Alina Striner

Instead of visualization she focused on "audiolization" - Mapping the number of traffic violations per hour to musical pitches in order to create an auditory representation of the data.

Table 1: Pitch, Midi pitch and number of accidents

Pitch	Midi pitch	Number of Accidents
A3	57	0-10
B3	59	11-20
C4	60	21-30
E4	64	31-40
G4	67	41-50
B6	69	51-60

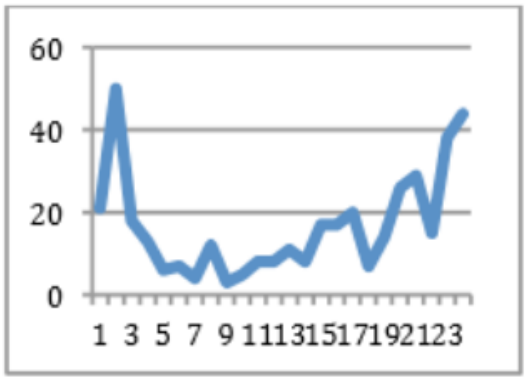


Figure 1: Visualization of occurrences per hour

Manjot Sawhney

Used data to simulate a baseball game.

First, allowed the user to place baseball players on different teams to create a league.

```
The teams are: Blue Jays, Orioles, and Nationals
Which team do you want to add Gregg Zaun to?
Jays
The teams are: Blue Jays, Orioles, and Nationals
Which team do you want to add Henry Blanco to?
orioles
The teams are: Blue Jays, Orioles, and Nationals
Which team do you want to add Moises Alou to?
nationals
The teams are: Blue Jays, Orioles, and Nationals
Which team do you want to add Corey Patterson to?
nationals
The teams are: Blue Jays, Orioles, and Nationals
Which team do you want to add Rod Barajas to?
Nationals

These are the players on the Blue Jays: Gregg Zaun.

These are the players on the Orioles: Henry Blanco.

These are the players on the Nationals: Moises Alou, Corey Patterson, Rod Barajas.
```

Created a baseball simulator which pitted two players against each other in 10 rounds of hitting. Used the players' batting averages, the pitchers' strike rates, and an element of Java randomness to determine whether the player hit the ball. Created a pop-up informing the user of the winner. Functional, user-friendly, polished, and included ideas for how to make a simulation for entire teams of players and to find the best hitter in the league.

```
It's time to play ball! In this game we have Rich Aurilia vs. Henry Blanco!

ROUND 1!

It's Rich Aurilia's turn to bat
It's a hit!
It's Henry Blanco's turn to bat
It's a hit!

The current score is
Rich Aurilia: 1
Henry Blanco: 1

ROUND 2!

It's Rich Aurilia's turn to bat
It's a miss!
It's Henry Blanco's turn to bat
It's a hit!

The current score is
Rich Aurilia: 1
Henry Blanco: 2

ROUND 3!

It's Rich Aurilia's turn to bat
It's a miss!
It's Henry Blanco's turn to bat
It's a miss!

The current score is
Rich Aurilia: 1
Henry Blanco: 2

ROUND 4!

It's Rich Aurilia's turn to bat
It's a miss!
It's Henry Blanco's turn to bat
It's a hit!
```