

CompSci 101

Introduction to Computer Science

APT: CompSci 101, Fall 2016, APT

This is the testing page. Once your program works here, you need to run your APT on the submit page (back on the previous page).

Problem Set 1	Details
First APTs are due on Sep 13, do them all	
<input type="radio"/> Bogsquare	
<input type="radio"/> Perimeter	
<input type="radio"/> Grayscale	
<input type="radio"/> BMI	together in class
<input type="radio"/> Gravity	do in lab 2
Test file: <input type="button" value="Browse..."/> No file selected.	
<input type="button" value="test/run"/>	

Sept 6, 2016

Prof. Rodger

Announcements

- Reading and RQ3 due next class
- Assignment 1 due today!
 - See the catch up schedule – for everyone!
- APT 1 due on Tuesday, Sep 13
- **Need a pin to add class** – fill out form
- Exam accommodations – fill out form
- Lab this week
 - try out functions, Solve APT
- Today – Problem Solving, Python practice, solve an APT

Algorithm

- Recipe
- Sequence of steps that constitute instructions
- Step-by-step procedure for calculations

What does Nate Silver do?

<http://53eig.ht/1tZy909>



How do Netflix and Amazon know me?

- Compsci101 project: capable of implementation as a program, but much more basic

<http://moreintelligentlife.com/content/features/anonymous/slaves-algorithm>

Who will win the presidency?



Chance of winning

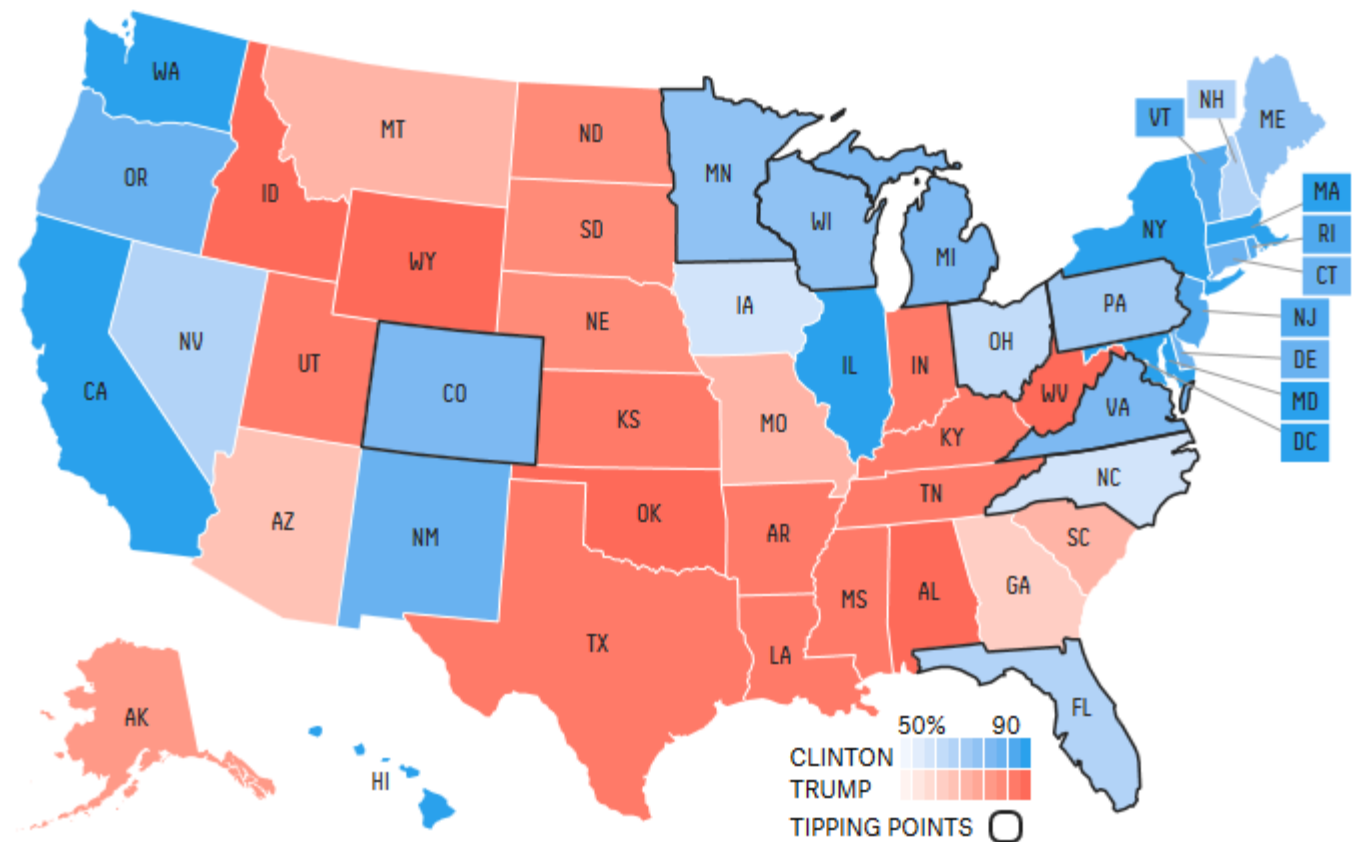


Hillary Clinton

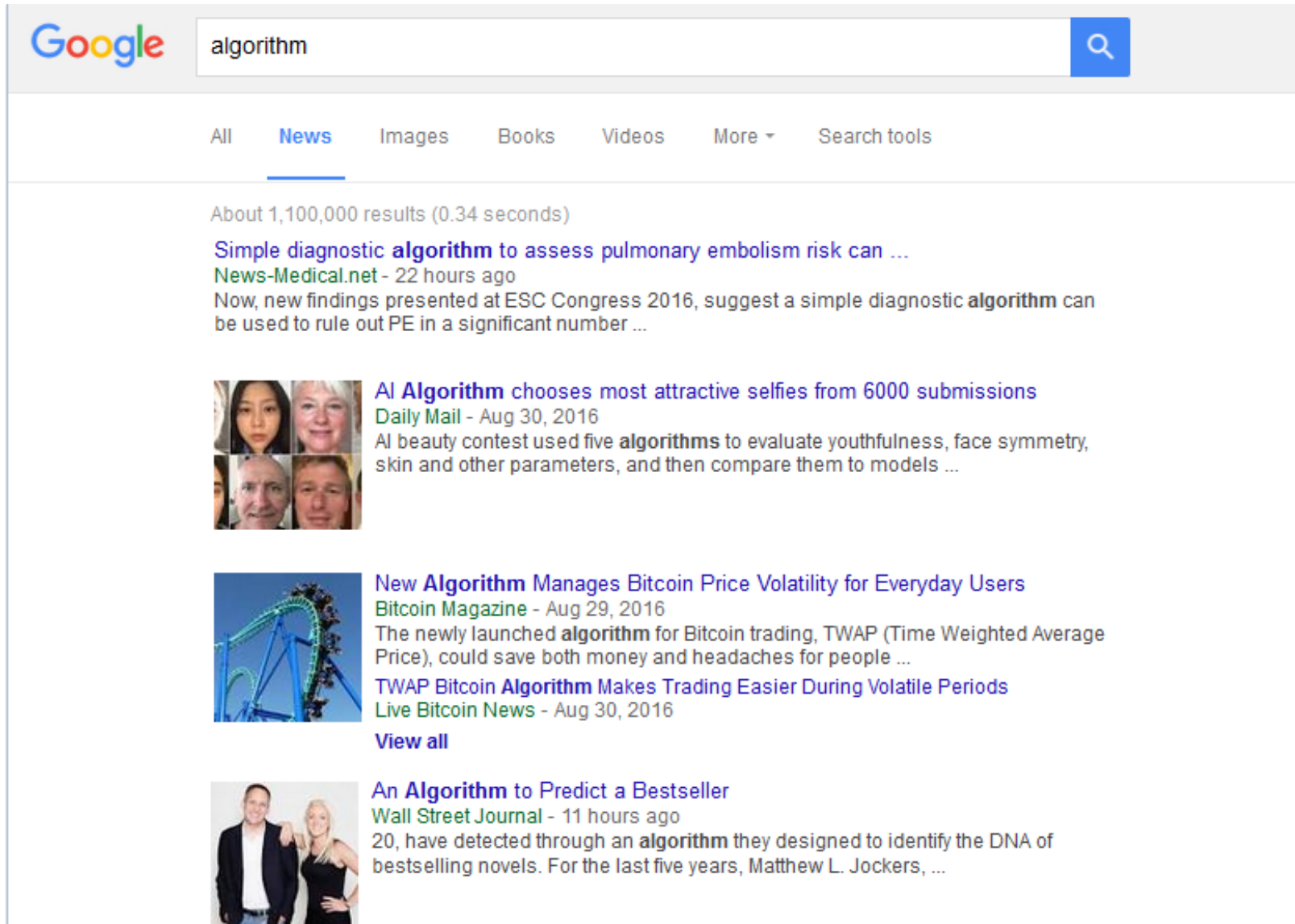
73.8%

Donald Trump

26.1%



Google “algorithm”



The image is a screenshot of a Google search results page for the query "algorithm". The Google logo is in the top left, and the search bar contains the word "algorithm" with a magnifying glass icon to its right. Below the search bar, there are tabs for "All", "News", "Images", "Books", "Videos", "More", and "Search tools". The "News" tab is selected and underlined. Below the tabs, it says "About 1,100,000 results (0.34 seconds)".

The first search result is titled "Simple diagnostic **algorithm** to assess pulmonary embolism risk can ...". It is from "News-Medical.net" and is dated "22 hours ago". The snippet reads: "Now, new findings presented at ESC Congress 2016, suggest a simple diagnostic **algorithm** can be used to rule out PE in a significant number ...".

The second search result is titled "AI **Algorithm** chooses most attractive selfies from 6000 submissions". It is from "Daily Mail" and is dated "Aug 30, 2016". The snippet reads: "AI beauty contest used five **algorithms** to evaluate youthfulness, face symmetry, skin and other parameters, and then compare them to models ...". To the left of the text is a small image showing four different faces.

The third search result is titled "New **Algorithm** Manages Bitcoin Price Volatility for Everyday Users". It is from "Bitcoin Magazine" and is dated "Aug 29, 2016". The snippet reads: "The newly launched **algorithm** for Bitcoin trading, TWAP (Time Weighted Average Price), could save both money and headaches for people ...". Below this, there is another line: "TWAP Bitcoin **Algorithm** Makes Trading Easier During Volatile Periods" from "Live Bitcoin News" dated "Aug 30, 2016". To the left of the text is a small image of a roller coaster. Below the snippet is a link that says "View all".

The fourth search result is titled "An **Algorithm** to Predict a Bestseller". It is from "Wall Street Journal" and is dated "11 hours ago". The snippet reads: "20, have detected through an **algorithm** they designed to identify the DNA of bestselling novels. For the last five years, Matthew L. Jockers, ...". To the left of the text is a small image of a man and a woman standing together.

Algorithms that scale: An example

- Human Genome Project
 - Multiple approaches, relying heavily on computational power and algorithms
 - Combine reads of DNA sequences, we'll look at an illustrative example
- These combine bio/chemistry techniques with computational techniques to recreate the sequencing, e.g., CGATTCCG... from "live data", actual DNA.

Eugene (Gene) Myers

- Lead computer scientist/software engineer at Celera Genomics, then at Berkeley, now at Janelia Farms Research Institute (HHMI)

"What really astounds me is the architecture of life. The system is extremely complex. It's like it was designed." ... "There's a huge intelligence there."

- BLAST and WG-Shotgun



Whole Genome Shotgun with words

olve problems.
ratively, create,
compsci101 we get t
01 we get to work colla
vely, create, and
s. In compsci1
y, create, and s
e get to work collabo
...

- Creation algorithm
 - Take a phrase
 - Replicate it four times
 - Chop into "chunks"
 - 15-22 characters
- How to recreate original phrase?

<http://bit.ly/101f16-0906-1>

From Algorithms to Code

- An algorithm that scales needs to run on a computer --- programming to the rescue!
- Extensive libraries help with programming
 - Brain or Neuroscience
 - Engineering and Mathematics
 - Genomics
 - Graphic User Interfaces, ...
- We are using Python, extensible and simple

Understanding terminology: code

- Move from "Hello World" to "Hello Around the World"
 - Look at Python, code, libraries
 - Learning (reviewing) terminology about Python

```
print "hello world"
```

```
f = open("hello.txt")  
for line in f:  
    print line
```

Hello around world code

<http://bit.ly/101f16-0906-2>

Running and Understanding Code

- Need Python compiler/interpreter
 - We're using Canopy, includes libraries
- Need an editor development environment
 - We use Eclipse and PyDev, open source and widely used, Ambient is Duke Plugin
- You need experience thinking and coding and debugging ideas and code:
 - Installing the suite of tools can be cumbersome
 - Persist, Persevere, Get Help, start over ☹

Barbara Liskov

- (one of) first women to earn PhD from compsci dept
 - Stanford 1968
- Turing award in 2008
 - Programming Languages



“It's much better to go for the thing that's exciting. But the question of how you know what's worth working on and what's not separates someone who's going to be really good at research and someone who's not. There's no prescription. It comes from your own intuition and judgment.”

Hello Around the World in Python

- We open a file, and we open a URL
 - Syntax slightly different, concept is similar
 - Real-world differences between files and URLs?

```
f = open("hello_unicode.txt")
```

```
f = urllib2.urlopen("http://nytimes.com")
```

- Must adhere to syntactic rules of Python
 - Naming, whitespace, : or . or (or) or [or]
- Must adhere to semantic rules of Python
 - Can't loop over anything, more rules to follow

Starting with Python

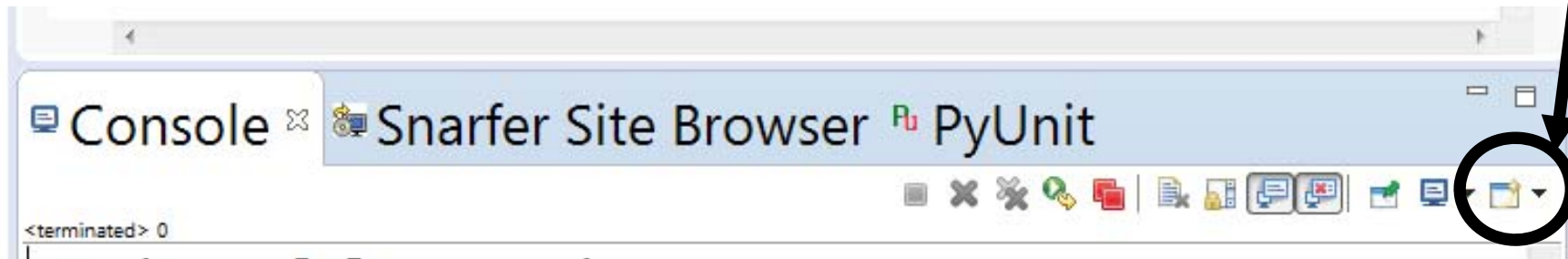
- Variable
 - Name of a storage location – holds a value
 - = to assign a value to a variable
- Type
 - Different types of data
 - A variable can stores data of a certain type
 - int, float, str
- operators in Python for numbers
 - Arithmetic: + - * / % **
- Built-in functions: pow, abs, round, int, float
example: `pow(2,3) + round (1.6)15`

Eclipse – Three ways to run

1. Write program and store in file
 - Create a PyDev project – a folder for programs
 - Create a PyDev module for each program (file)
 - Run in console
2. Create an APT in Eclipse and run on web
3. Run interactively
 - Open PyDev console
 - Execute each line as typed
 - Code not saved

Demo: Run interactively in Eclipse PyDev Console

- If Console window is not showing then
 - Click on Window, Show View, Console
- Then at the bottom of Eclipse, click here:



- Select PyDev Console, Python Console

Variables, Types, Expressions?

```
a = 5
```

```
b = 4
```

```
print b
```

```
a = a + b
```

```
print a
```

```
c = "fred"
```

```
print c
```

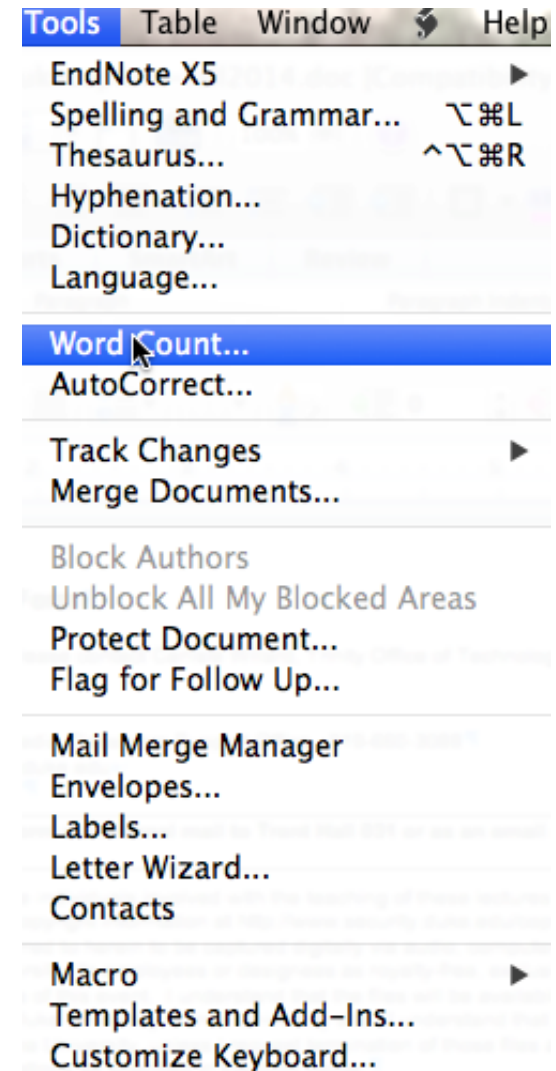
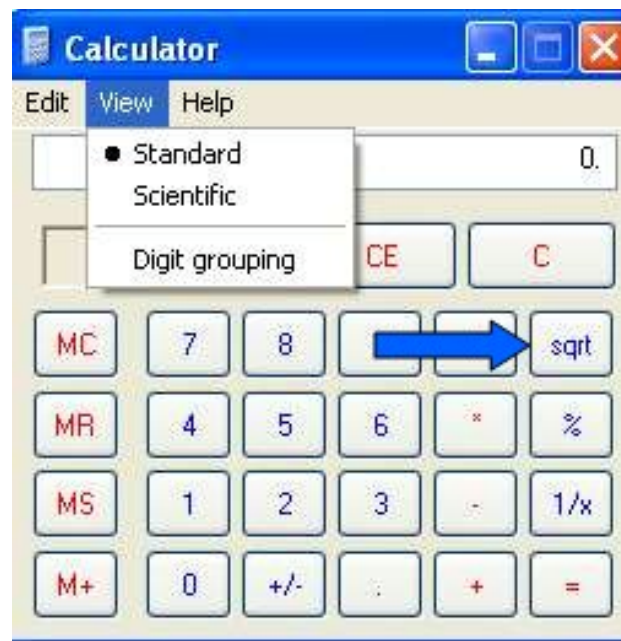
```
print a + b * 3
```

```
print (a + b) *  
3
```

```
print a / b
```

```
print a / (b *  
1.0)
```

Examples of functions



Functions explained

- In a calculator, sqrt: number in \rightarrow number out
 - What is domain, what is range?
- In MSWord, word count: document \rightarrow number
 - Domain is word doc, range is integer
- In browser, web: URL \rightarrow HTML formatted "page"
 - Domain is valid URL, range is HTML resources
- In Python we see similar structure!

Demo

- In Eclipse write a file with a function and run it
- stuff.py

```
def sum(a, b):  
    return a+b  
  
print sum(3,5)  
print sum(1,4)
```

Python Functions

- Answer these questions based on thinking, don't run any code
 - <http://bit.ly/101f16-0906-3>
- Why do we need functions?
 - Manage complexity of large programs
 - Test and develop code independently
 - Reuse code in new contexts: create APIs!

Functions return values

- Most functions return values
 - Sometimes used to make things simpler, but returning values is a good idea

```
def inch2centi(inches):  
    return 2.54*inches
```

```
xh = inch2centi(72)
```

```
def pluralize(word):  
    return word + "es"
```

```
pf = pluralize("fish")
```

Functions can print info

- Some functions only print info
- Note there is no return statement in the function

```
def helloPerson(name):  
    print "hello" + name  
  
helloPerson("Susan")  
helloPerson("Ademola")
```

APTs

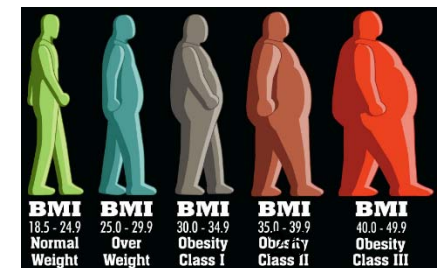
APT: CompSci 101, Fall 2016, APT

This is the testing page. Once your program works here, you need to run your APT on the submit page (back on the previous page).

Problem Set 1	Details
First APTs are due on Sep 13, do them all	
<input type="radio"/> Bogsquare	
<input type="radio"/> Perimeter	
<input type="radio"/> Grayscale	
<input type="radio"/> BMI	together in class
<input type="radio"/> Gravity	do in lab 2
Test file: <input type="button" value="Browse..."/> No file selected.	
<input type="button" value="test/run"/>	

What is an APT? BMI APT

- Automated/Algorithmic Problem Testing
 - Write one function, 2-30 lines, solve a problem
 - Tested automagically in Eclipse or the browser
 - Lots of test cases – test test test
- Start simple, build toward more complex
 - What is a function? A function call?
 - What is a parameter? Argument?
 - How do you run/execute a program



Demo Solving APT BMI

- Write your code in Eclipse
 - Create python file
 - Name of file important – case matters
 - name of function important – cut and paste this
 - Write your code
 - Test a few examples in Eclipse
- Run online on using APT Tester
 - Tests on lots of examples, Debug, fix
 - Get all **GREEN**
- Submit on APT page
 - README form too