**Are the Millennials Engaged?**

You work at the global megacorp social network SpyPy. SpyPy has 1.5 billion daily users, and you want to make sure that people in the millennial age bracket are engaging with your website. Your boss seems particularly frazzled by this question, and he's put it on you to find out. You decide that "engagement" means spending more than the average of seven minutes on the website. You fire up your data-science stack in Python and first check the average time -- which turns out to be near 11 whole minutes! But you can't really tell if they're *really* spending more time or if it's just random chance that a few of your users left the browser open and walked away. You write the following code:

In [5]:

**import** **spypy**

**from** **scipy.stats** **import** ttest\_1samp

millennial\_times = spypy.get\_site\_times\_for\_demographic('millennial')

t\_stat, p\_val = ttest\_1samp(millennial\_times, 7)

**if** p\_val < .05:

**print** "The Millennials are engaged!"

**else**:

**print** "The Millennials are not engaged :(!"

The Millennials are engaged!

**SpyPy: We're Significantly Different**

Well that's great news! Millennials are, for the most part, spending around 10 minutes on your website. But before you break out the champagne glasses your boss is in a frenzy again, this time about Metropolitan Statistical Areas (MSAs). You are tasked with finding if people in cooler climates post more pictures on SpyPy than people in warmer climates. You cross corroborate with weather data and run a statistical test on the info.

In [2]:

**from** **scipy.stats** **import** ttest\_ind

warmer\_weather\_picture\_count = spypy.get\_number\_pictures\_for\_climate('hot')

colder\_weather\_picture\_count = spypy.get\_number\_pictures\_for\_climate('cold')

t\_stat, p\_val = ttest\_ind(warmer\_weather\_picture\_count, colder\_weather\_picture\_count)

**if** p\_val < .05:

**print** "People from colder climates post a different number of pictures compared to people from warmer climates"

**else**:

**print** "Climate doesn't appear to affect the number of pictures posted"

Climate doesn't appear to affect the number of pictures posted

**SpyPy: Because We Care About Your Data**

Seems like climate *probably* doesn't really affect the number of times people post pictures. Not really sure why that would've been the case anyway. SpyPy has a new feature that you think will get people to interact with the website for longer: SpyPy Stories. It is preliminarily being launched to 8 million users and the internal goal is to get 2 million people to post SpyPy Stories in the first week. Unfortunately, only 1,997,893 people posted SpyPy Stories this week. We want to know if this is a significant difference from our goal -- did we pretty much meet it or did we seriously miss? You know how to answer this question:

In [7]:

**from** **scipy.stats** **import** binom\_test

number\_of\_trials = 8000000

expected\_successes = 2000000

actual\_successes = 1997893

expected\_success\_rate = float(expected\_successes) / float(number\_of\_trials)

p\_val = binom\_test(actual\_successes, n=number\_of\_trials, p=expected\_success\_rate)

**if** p\_val < 0.05:

**print** "We didn't hit our target by a significant amount"

**else**:

**print** "We just missed our target by a very small amount!"

We just missed our target by a very small amount!

Looks like we came very close to hitting our target for SpyPy Stories! You've saved the day so many times already! Your boss comes by to thank you for all the hard work you put in today and says you've made significant contributions to the team. You tell her you're not sure if that's true, but you definitely have a way of finding out.