ORGANIZING THE GLOBAL HISTORICAL CLIMATOLOGY NETWORK

JILL HARDY

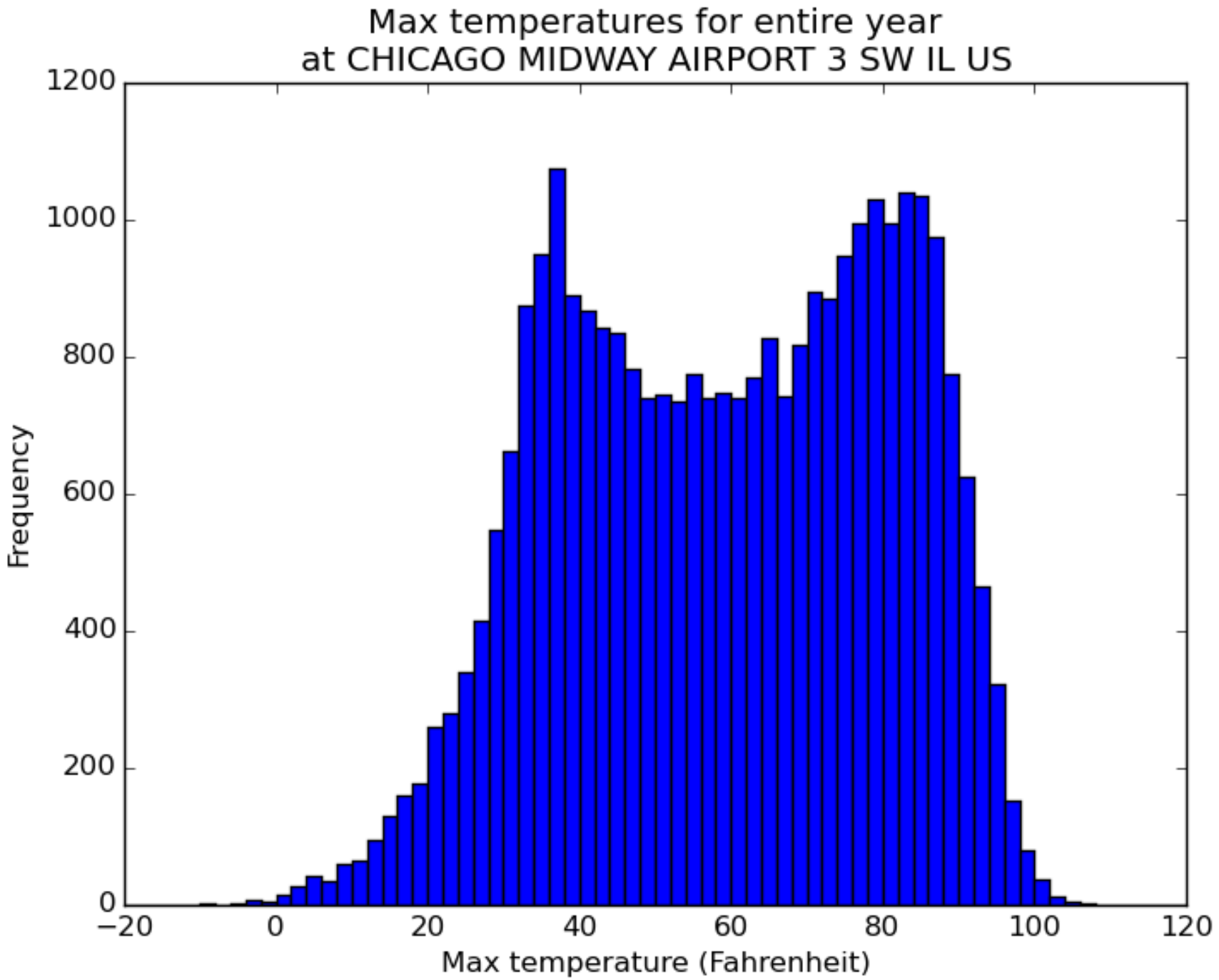
RACE CLARK

CHRIS NATOLI

WILLIAM MATTHEWS

TEMPERATURES: AVERAGES, RECORDS, AND EVERYTHING IN BETWEEN

- What is “normal”?
- What is a “record”?
- “Strong” and “weak” records
- Climate change



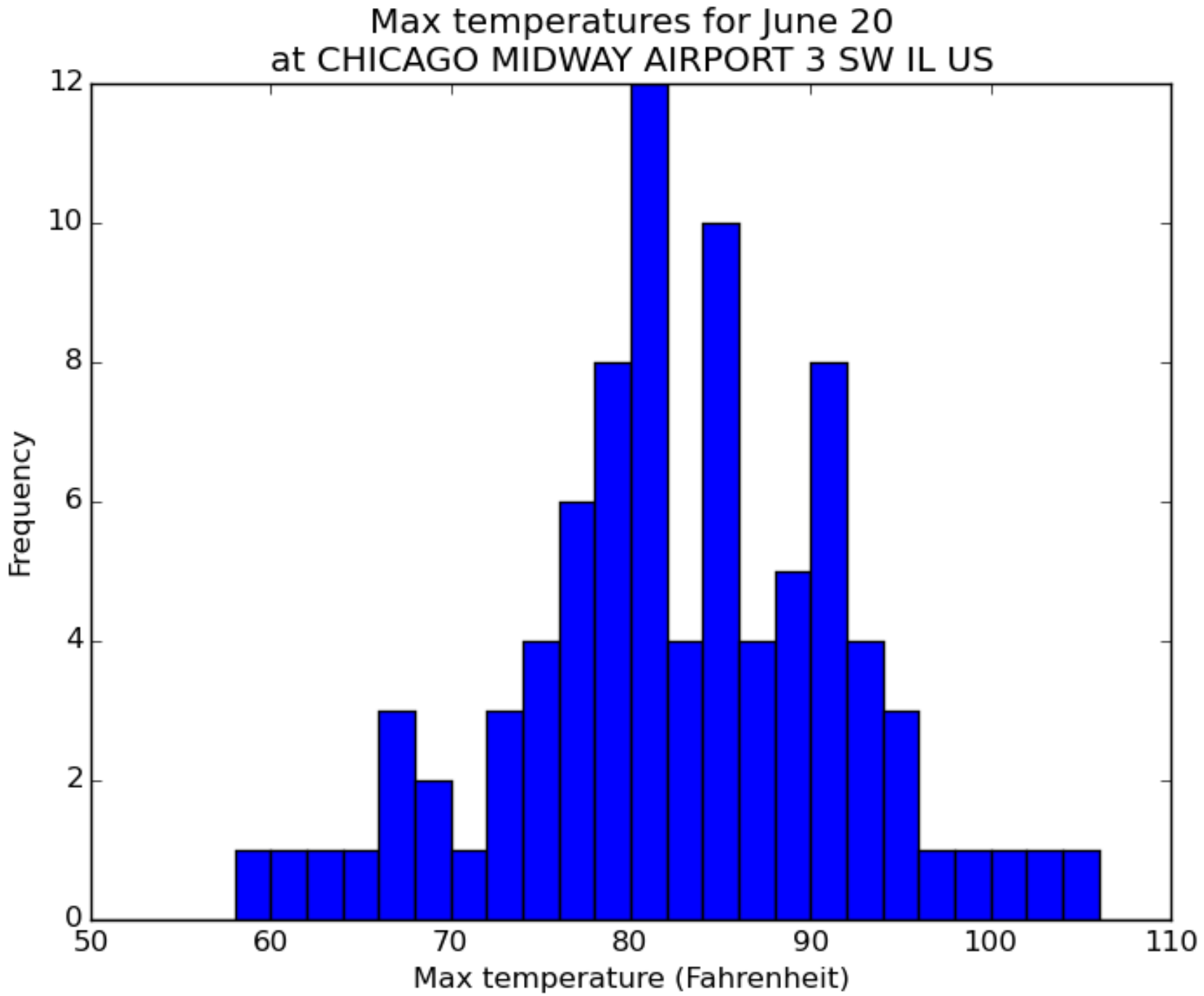
86 years of maximum temperatures

Over 31,000 total measurements

Bimodal

Wide range – continental climate

Can tell us rarity of all-time highs or all-time lows



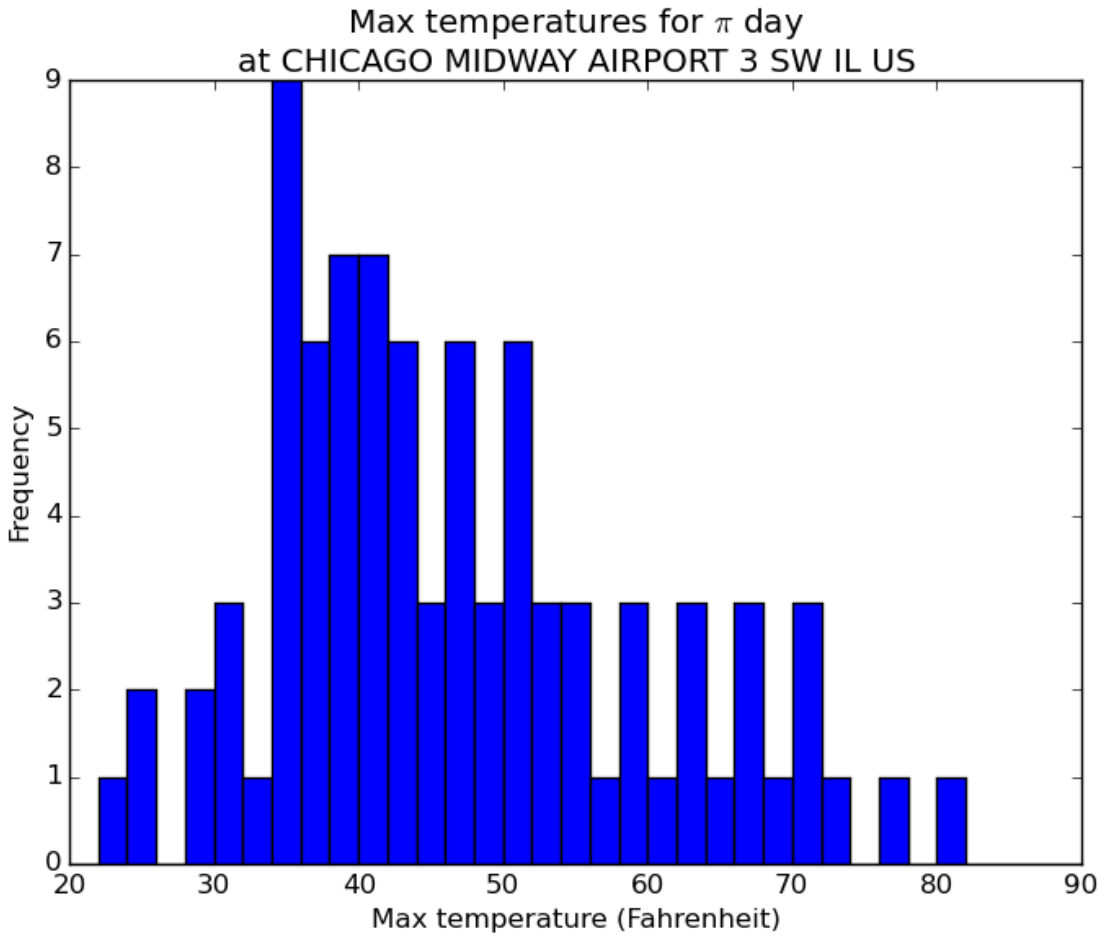
Now only 86 data points

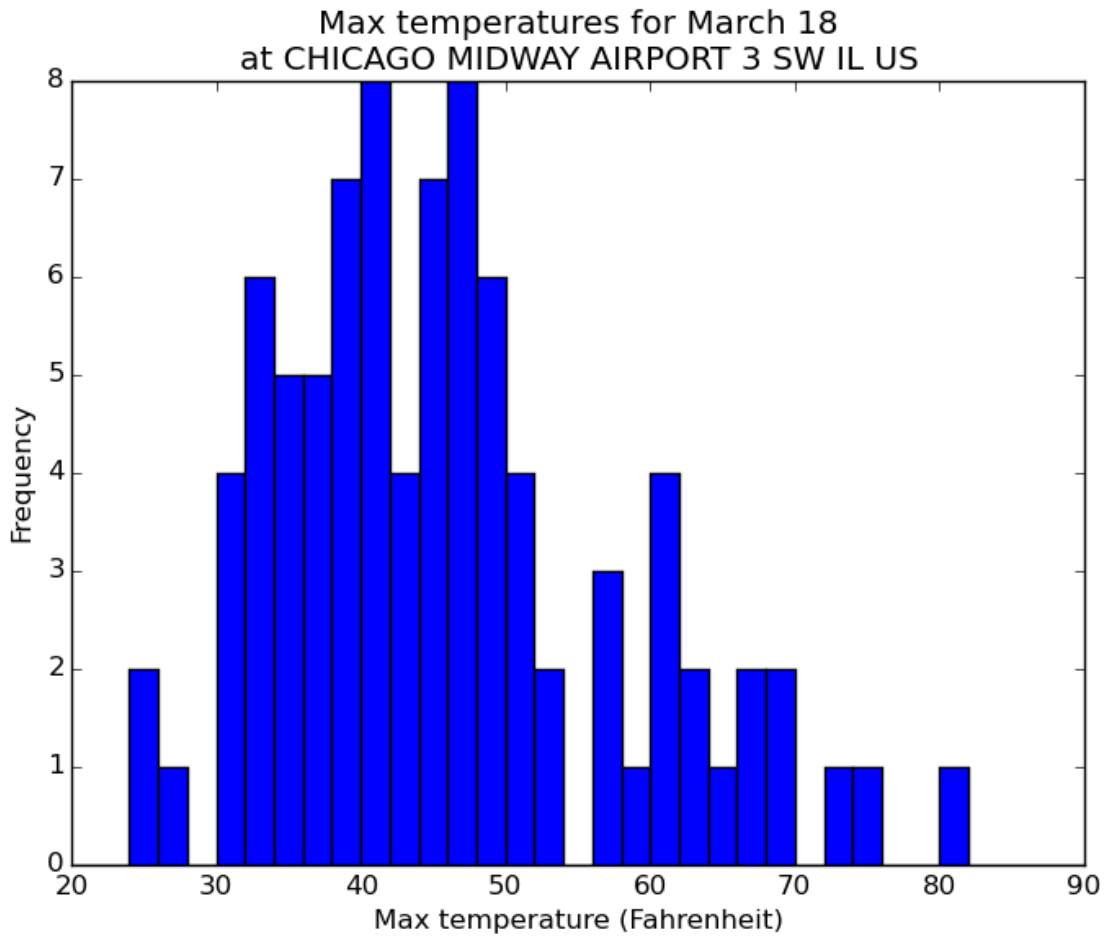
Normal distribution

Wide variation still

No outliers

What if we could plot the temperature ranges as an animation through the year?





Outliers! (20s on Pi Day and Spring Break …. Or 80s)

Wide variation

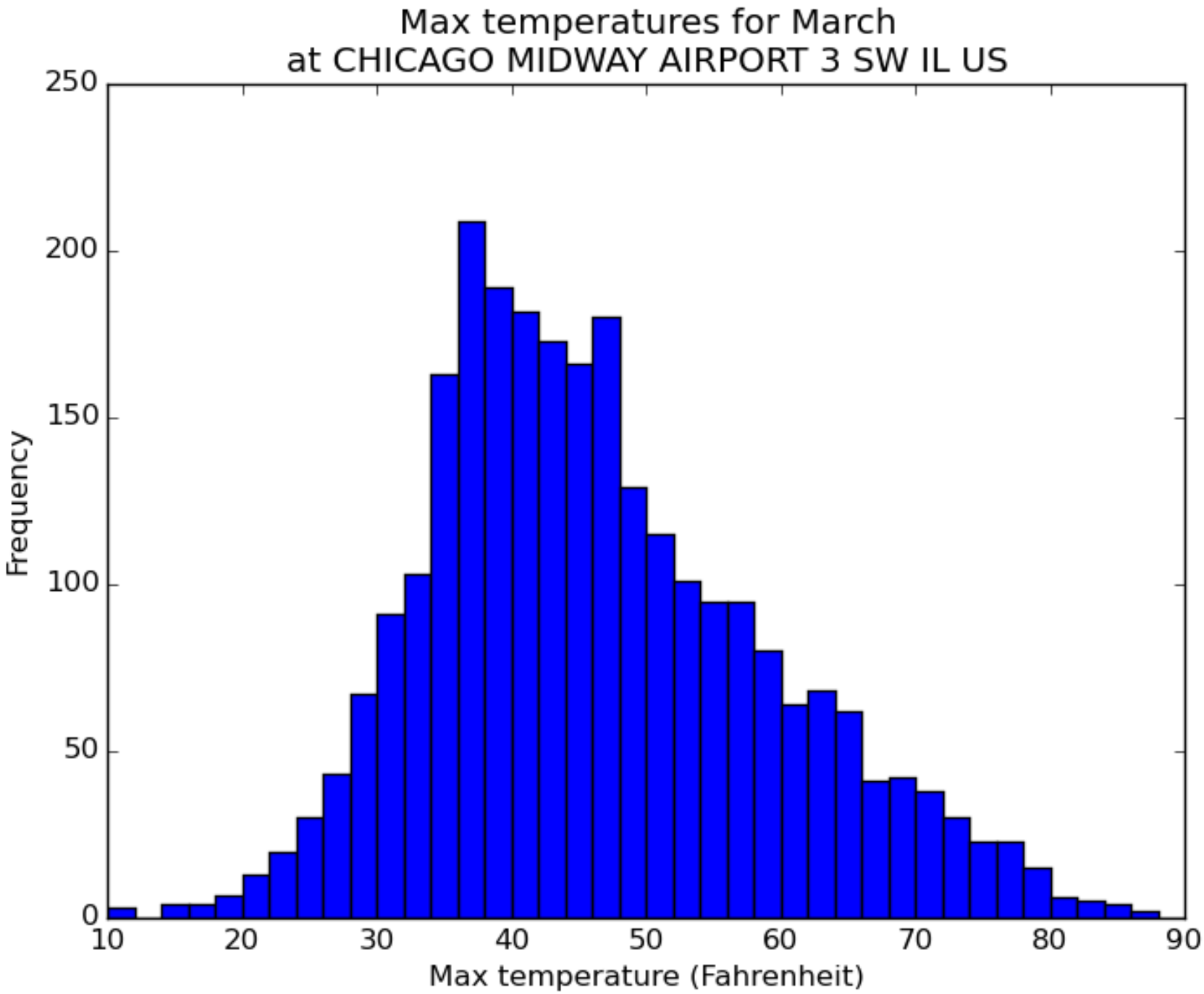
Long warm tail

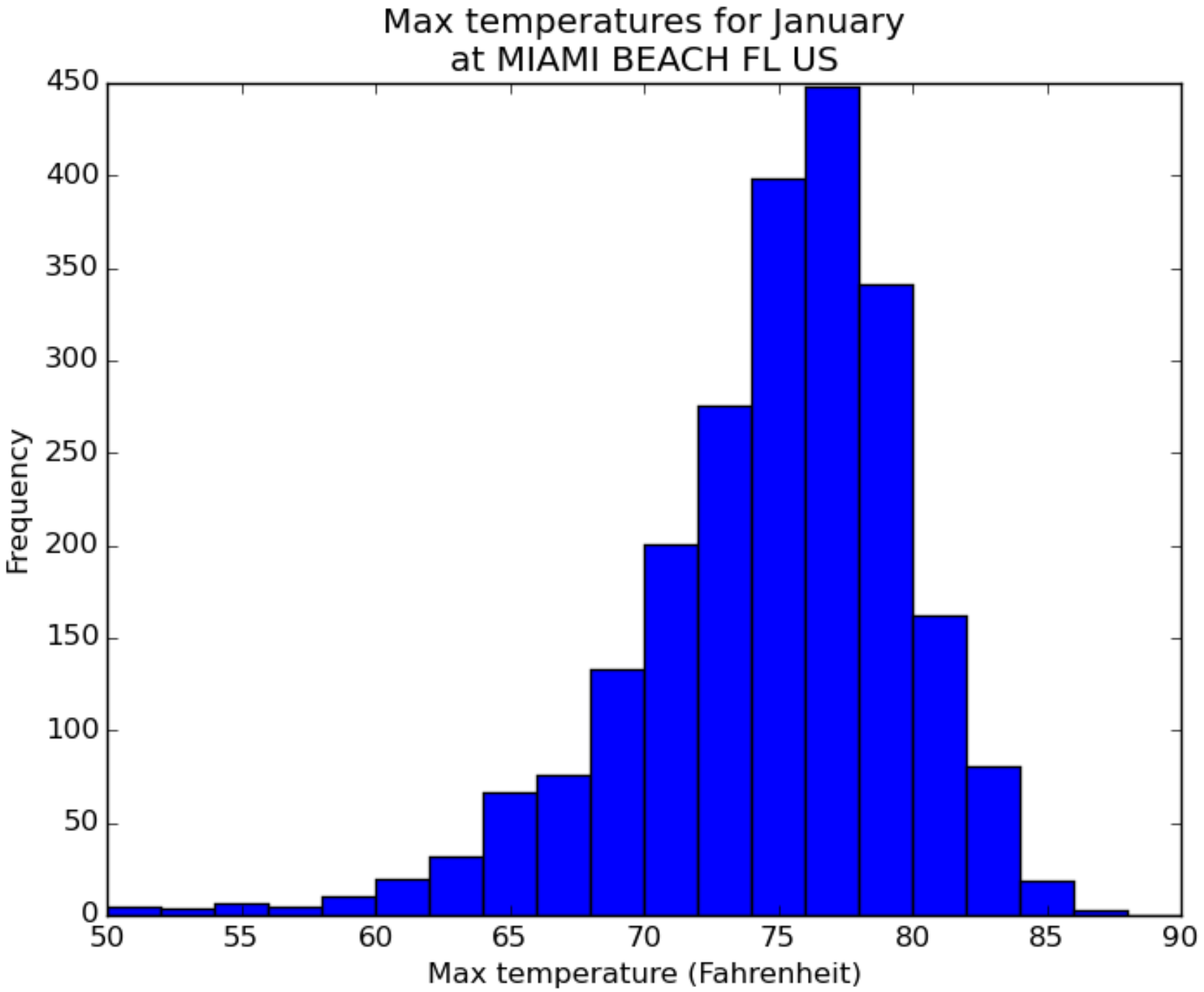
Nearly 2,700 points

All those above 80 recorded in one week in 2012

Extremely unlikely!

Note 10 to 90F range….

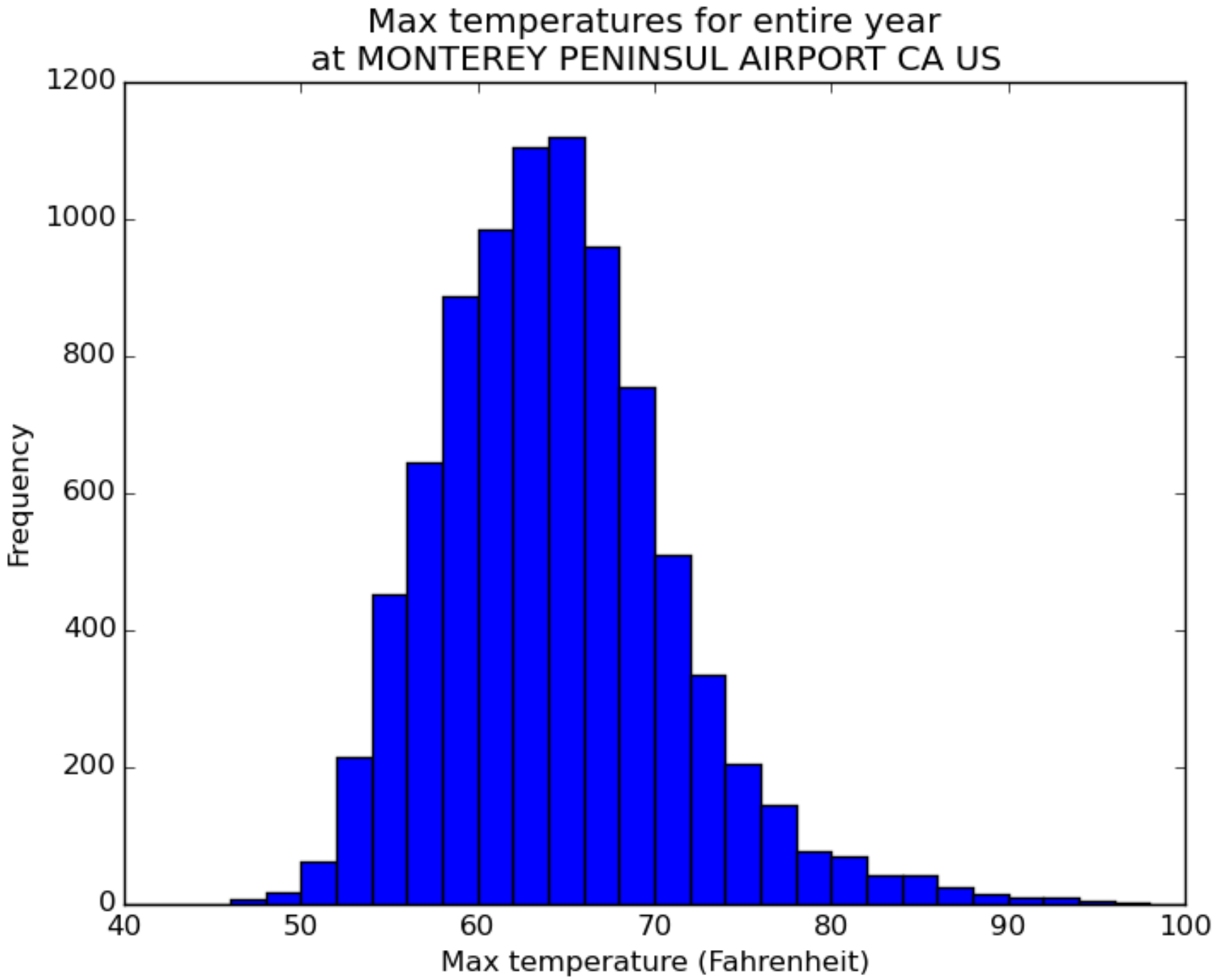




Tighter distribution

Subtropical climate

Now you see a cold tail



Tighter distribution

Mediterranean climate

When people in California say 80 or 90 is hot, you know they mean it….

```python
import matplotlib.pylab as plt
import csv

# Convert temperature from tenths of degree Celsius to degree Fahrenheit
# and round to nearest int.
def convert_temp(c):
    c = float(c) / 10
    f = 9 / 5 * c + 32
    return int(f + 0.5)



num_to_month = {1: 'January', 2: 'February', 3: 'March', 4: 'April',
                5: 'May', 6: 'June', 7: 'July', 8: 'August',
                9: 'September', 10: 'October', 11: 'November', 12: 'December'}
filenames = ('chicago', 'chula_vista', 'miami', 'memphis', 'monterey', 'okc',
             'vienna')


# For each location, plot a histogram of the entire year of data
# and one for each month.
for filename in filenames:
    data = []
    with open(filename + '.csv') as input_file:
        reader = csv.reader(input_file)
        next(reader)
        for row in reader:
            for i in range(1,3):
                row[-i] = convert_temp(row[-i])
            data.append(row)

    # Extract all max temperatures and monthly max temperatures.
    tmaxes = [ int(row[-2]) for row in data if row[-2] != -999.9]
    monthly_tmaxes = [ [ row[-2] for row in data
                         if row[-2] != -999.9 and int(row[-3][4:6]) == i ]
                       for i in range(1,13) ]

    # Plot a histogram of the entire year of data. Bin width is 2 degrees.
    plt.hist(tmaxes, 100, (-50,150))
    plt.xlabel('Max temperature (Fahrenheit)')
    plt.ylabel('Frequency')
    plt.title('Max temperatures for entire year\nat {}'.format(data[0][1]))
    plt.savefig('plots/' + filename + '_all.png')
    plt.close()

    # Plot a histogram for each month with bin width of 2 degrees.
    for i in range(1,13):
        plt.hist(monthly_tmaxes[i-1], 100, (-50,150))
        plt.xlabel('Max temperature (Fahrenheit)')
        plt.ylabel('Frequency')
        plt.title('Max temperatures for {}\nat {}'.format(num_to_month[i],
                                                          data[0][1]))
        plt.savefig('plots/' + filename + '_' + num_to_month[i] + '.png')
        plt.close()

# Also, plot March 18 data for Chicago since it was a particularly
# extreme day in 2012.
data = []
with open('chicago.csv') as input_file:
    reader = csv.reader(input_file)
    next(reader)
    for row in reader:
        for i in range(1,3):
            try:
                row[-i] = convert_temp(row[-i])
            except:
                print(row)
        data.append(row)
day_tmaxes = [ int(row[-2]) for row in data
               if row[-2] != -999.9 and row[-3][-4:] == '0318' ]
plt.hist(day_tmaxes, 100, (-50,150))
plt.xlabel('Max temperature (Fahrenheit)')
plt.ylabel('Frequency')
plt.title('Max temperatures for March 18\nat {}'.format(data[0][1]))
plt.savefig('plots/chicago_march18.png')
```