**Name: Mehmet Nuri Akyol**

**Student Number: 213910467**

**Hypothesis: Between 1880 to 2000 years teams avarage wins are have difference to 2000 to 2015 years. Between 2000 to 2015 years teams have very similar wins rate and I want to proof these are correct.**

**Codes:**

**Firstly, I import my working codes which are :**

import thinkstats2

import thinkplot

import pandas as pd

import numpy as np

teams = pd.read\_csv("Teams.csv")

from \_\_future\_\_ import print\_function, division

%matplotlib inline

**After that I have to divide years and I have to calculate teams wins to years and I did these with this codes:**

\_1880to\_2000 = teams[teams.yearID < 2000]

win\_to\_years=\_1880to\_2000.W

win\_to\_years

**And I measured avarage wins relation with years :**

year\_win\_rate=dict()

for i in range(1880,2000):

year\_win\_rate[i] = teams.W[teams.yearID == i].mean()

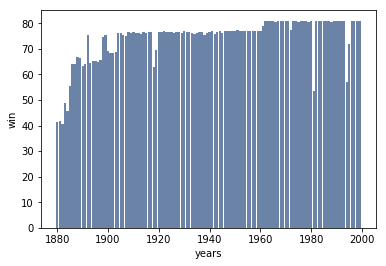
year\_win\_rate

**After these codes I plot the graphs with pmf, cdf and hist with these codes:**

year\_win\_rate\_graph = thinkstats2.Hist(year\_win\_rate)

thinkplot.Hist(year\_win\_rate\_graph)

thinkplot.Config(xlabel="years", ylabel="win")



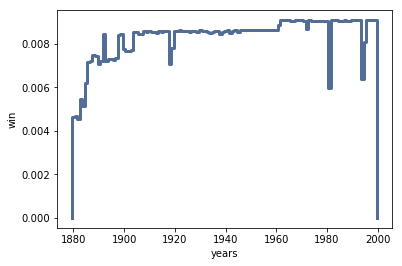
This graph is working with hist.

And I work with pmf and cdf :

year\_win\_rate\_graphpmf = thinkstats2.Pmf(year\_win\_rate)

thinkplot.Pmf(year\_win\_rate\_graphpmf)

thinkplot.Config(xlabel="years", ylabel="win")

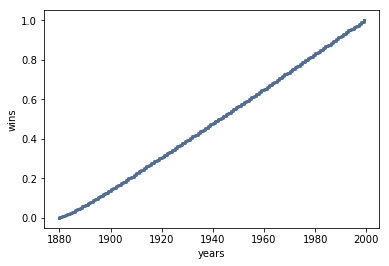


After pmf, I use Cdf graphs with this code and grahps:

year\_win\_rate\_graphcdf = thinkstats2.Cdf(year\_win\_rate)

thinkplot.Cdf(year\_win\_rate\_graphcdf)

thinkplot.Config(xlabel="years", ylabel="wins")



**And this point I finished first part of project I measured 1880 to 2000 years teams wins rate and after this point I want to proof 2000 to 2015 years team wins and difference between early years**.

\_2000to\_2017 = teams[teams.yearID > 2000]

win\_to\_yearsup=\_2000to\_2017.W

win\_to\_yearsup

year\_win\_rateUp=dict()

for i in range(2000,2017):

year\_win\_rateUp[i] = teams.W[teams.yearID == i].mean()

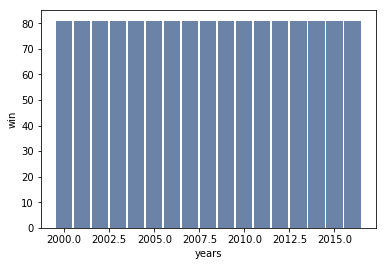
year\_win\_rateUp

**I will divide years again. And after this I will plot hist, pmf and cdf graphs for this datas.**

year\_win\_rateUp\_graph = thinkstats2.Hist(year\_win\_rateUp)

thinkplot.Hist(year\_win\_rateUp\_graph)

thinkplot.Config(xlabel="years", ylabel="win")



**And after hist , I will plot pmf and cdf graphs:**

year\_win\_rateUp\_graphpmf = thinkstats2.Pmf(year\_win\_rateUp)

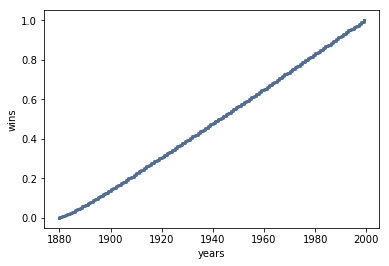
thinkplot.Pmf(year\_win\_rateUp\_graphpmf)

thinkplot.Config(xlabel="years", ylabel="win")

**And I have code for cdf graphs with these codes and I plotted cdf graphs:**

year\_win\_rateUp\_graphcdf = thinkstats2.Cdf(year\_win\_rate)

thinkplot.Cdf(year\_win\_rateUp\_graphcdf)

thinkplot.Config(xlabel="years", ylabel="wins")

**As a result, from these graphs, it is clear that the games won by teams until 1880 to 2000 are more irregular than 2000 to 2015 teams won games.**