HISTOGRAM

November 5, 2015

In [1]:
```
import matplotlib.pyplot as plt
import csv
import math
import numpy as np
import matplotlib
```

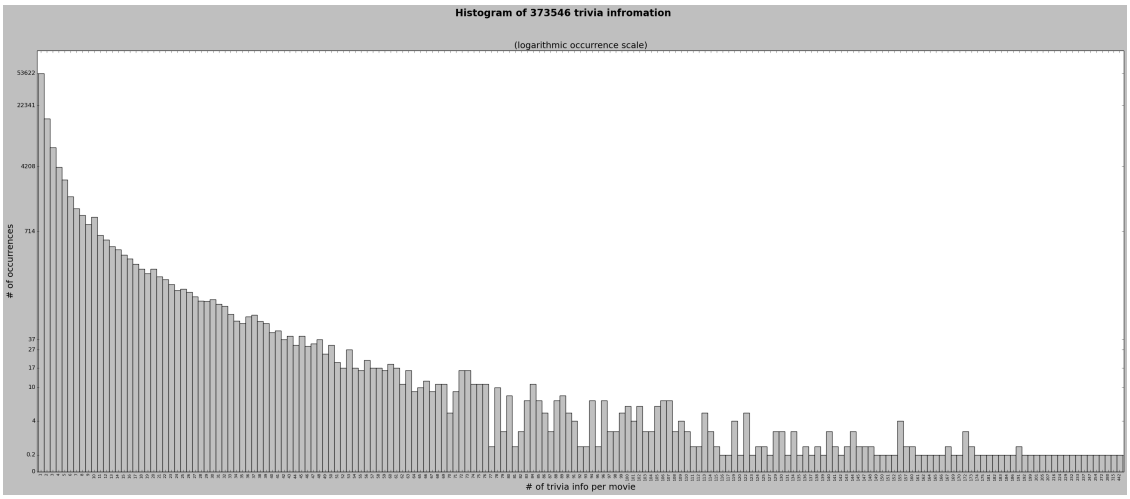
In [2]:
```
a = []
with open('count_of_count_of_triva.csv','r') as f:
    reader = csv.reader(f)
    for row in reader:
        a.append(row)
b =[]
for i in a:
    if int(i[1])==1:
        b.append([i[0],math.log10(int(i[1]))+0.2])
    else:
        b.append([i[0],math.log10(int(i[1]))])
b = np.asarray(b)
a = np.asarray(a)
#print a
a_y = a[:,1]
y = []
y = a_y.tolist()
y = [int(i) for i in y]
for i in y:
    if i ==1:
        i ==2

#print a
#print a_y
print max(a_y)
print min(a_y)
b_count = b[:,0]
b_log = b[:,1]
g = []
g = b_log.tolist()
#print max(g)
m = []
m = [float(i) for i in g]
#print max(m)
x_pos = np.arange(len(b_count))
fig_size = plt.rcParams["figure.figsize"]
print "Current size:", fig_size
fig_size[0] = 40
```

```
        fig_size[1] = 15
        plt.rcParams["figure.figsize"] = fig_size
        print len(b_log)
        low = min(m)
        high = max(m)
        def number(x,pos):
            if x ==0:
                return int(x)
            if x == 0.2:
                return x
            else:
                return int(math.pow(10,x))
        print "max m",max(m)
        print "min m",min(m)
        print "max y",max(y)
```

```
98
1
Current size: [8.0, 6.0]
183
max m 4.72934300832
min m 0.2
max y 53622
```

```
In [7]: import matplotlib.ticker
        from matplotlib.ticker import FuncFormatter
        width = 1.0
        fig = plt.figure()
        fig.suptitle('Histogram of 373546 trivia infromation', fontsize=20, fontweight='bold')
        ax = plt.axes()
        ax.set_title('(logarithmic occurrence scale)',fontsize = 18)
        ax.set_xlabel('# of trivia info per movie',fontsize = 18)
        ax.set_ylabel('# of occurrences',fontsize = 18)
        #ax.set_yscale('log')
        ax.set_yticks([0,0.2,1,0.602059991328,1.23044892138,1.44715803134,1.56820172407,2.85410843915,3
        ax.set_xticks(x_pos + (width / 2))
        ax.set_xticklabels(b_count,rotation=90)
        ax.tick_params(axis='x', labelsize=8)
        #ax.get_yaxis().get_major_formatter().labelOnlyBase = False
        #ax.yaxis.set_major_formatter(matplotlib.ticker.FuncFormatter(lambda x, pos: str(int(round(x))))
        plt.bar(x_pos,m,width,color = 'silver')
        plt.ylim([0, math.ceil(high+0.1)])
        formatter = FuncFormatter(number)
        ax.yaxis.set_major_formatter(formatter)
        #plt.yscale('log', nonposy='clip')
        plt.margins(0.001)
```



In [203]:

```
<matplotlib.figure.Figure at 0xb12c082c>
```

In [ ]: