**Total Seek Count For Fall 2013**

Goal: Calculation of total seeks for each video for Fall 2013

The play count program that was used to calculate the total play events in a video was modified.

For the event type instead of “play\_video”, it was changed to “seek\_video”

Every time a user seeks a video from one spot to another, an event is created named as “seek\_video”. Each “seek\_video “event gives two sets of time, one is “old\_time” and the other is “new\_time”. “Old time” gives the exact time point during the video when the video was scrubbed while “new\_time” gives the time point in the video where the video was scrubbed to. We are paying more attention to the “new time” because this will help us identify the sections that student were more engaged in. This “new time” may suggest interesting/ confusing topics or something that was a major part of the exam.

The program is shown below:

videoNames = {}

with open("Video Names 3.csv", "rU") as f:

reader = csv.reader(f)

for row in reader:

videoNames[row[0]] = row[1] #cannot remember how this line works exactly

usernames = Set()

with open("Names for SPOC.csv", "rb") as f:

reader = csv.reader(f)

for row in reader:

usernames.add(row[3])

data={} #first dict with username (key) : second dict (value)

def parseEventText(eventLine):

line = eventLine.replace("{","")

line = line.replace("}","")

units = line.split(",")

result = {}

for unit in units:

pieces = unit.split(":")

result[pieces[0].replace('"',"")] = pieces[1].replace('"',"")

return result

for line in liblytics.read\_log\_file("tracking\_700x\_UMass\_\_Fall\_2013.log.gz"): #Reads line in log file

if (line["event\_type"] == "seek\_video"): # Grabs only seek\_videos

username=line["username"]

newtimes = parseEventText(line["event"])["new\_time"]

#first loop creates first dict

if username != "":

videoName = parseEventText(line["event"])["id"]

if username not in data:

data[username] = {}

usersDict = data[username]

#creates third dict with playCount and Times (keys) and their values

if videoName not in usersDict: #second loop creates third dict

videoDict = {}

videoDict["New Times"] = []

videoDict["Seek Video Count"] = 1

usersDict[videoName] = videoDict

videoDict = usersDict[videoName]

videoDict["New Times"].append(newtimes)

videoDict["Seek Video Count"] = len(videoDict["New Times"])

usersDict[videoName] = videoDict

data[username] = usersDict

# Prints to CSV

f = open("Total Seek Videos f13.csv", "w") #This is where you could name the CSV

f.write("username,")

for cleanName in sorted(videoNames):

f.write(cleanName)

f.write(",")

f.write(" \n")

##for username in sorted(data):

##if username in usernames:

for username in sorted(usernames):

if username in data:

f.write(username)

f.write(",")

for cleanName in sorted(videoNames):

if videoNames[cleanName] in data[username]:

f.write(str(data[username][videoNames[cleanName]]["Seek Video Count"]))

f.write(",")

else:

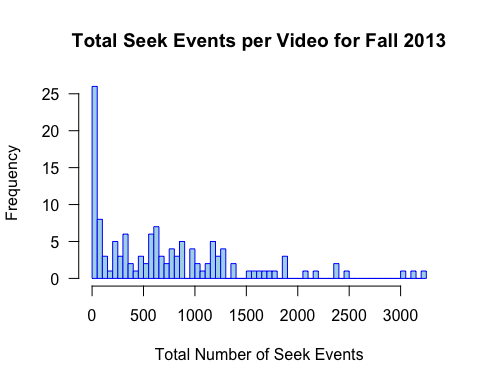
f.write("0,")

f.write(" \n")

f.close()

The program currently gives the total number of “seek\_video” events for each videos of Fall 2013.

A histogram was generated of total number of seek events for all videos. Totals are put in bins with a difference of 100 seek events.



The videos with total number of seek events less than 1 standard deviation (711.5726) below the mean were listed as the least seeked videos while the videos with more than 1 standard deviation above the mean (718.2769) were listed as the most seeked videos.

Least seeked videos :Total <7

Highest seeked videos: Total>1430

**Top 10 Most seeked videos (F13) : Total>1430) – More than 1**

**standard deviation above the mean ( Arranged in descending order)**

1) Name of the video -06.3-AminoAcids Number of seeks- 3230

Video by Eric Lander Length of the video: 17:49

From the topic Protein structure

Followed by 1 Test yourself question

2) Name of the video - Lab18.1-YeastInTheLab Number of seeks-

3144

Length of the video: 6:32

This is a lab video that talks about the budding yeast. Video shows

growing yeast colonies in a petri dish.

3) Name of the video -25.4-PCR Number of seeks- 3042

Video by Eric Lander Length of the video: 14:16

It is part of the Recombinant DNA 2 section. Video explains the

process of polymerase chain reaction. It is followed by one test

yourself question.

Talk about primer matching- explain 5’ and 3’ ends.

How is PCR useful and some examples where PCR is used.

4) Name of the video - CuttingAndPastingDNA Number of seeks-

2458

Video by Eric Lander Length of the video: 21:42

From the topic Recombinant DNA 1

Followed by 2 Test yourself questions

5) Name of the video - 21.1-ReplicationInDifferentOrganisms

Number of seeks- 2373

Video by Eric Lander Length of the video: 28:43

Followed by 1 Test yourself question

From the topic Introns

6) Name of the video - 22.2-BetaGlobinMutations Number of seeks-

2369

Video by Eric Lander Length of the video: 22:34

Followed by 1 Test yourself question

From the topic Mutations

7) Name of the video - 06.2-PrimaryStructureOfProteins Number of

seeks- 2184

Video by Eric Lander Length of the video: 10:08

Followed by 1 Test yourself question

From the topic Protein Structure

8) Name of the video - 04.1-FractionatingLife Number of seeks-

2083

Video by Eric Lander Length of the video: 14:39

Followed by 1 Test yourself question

From the topic Hydrogen Bonds

9) Name of the video - 10.1-EnergeticsOfPathways Number of seeks-

1896

Video by Eric Lander Length of the video: 06:39

Followed by 1 Test yourself question

From the topic Glycolysis

10) Name of the video -26.1-TheGenomicLandscape Number of

seeks- 1879

Video by Eric Lander Length of the video: 21:15

The video is part of the genomics section. It is followed by six test

yourself questions. Video talks about DNA, RNA, process of

transcription, reverse transcriptase and briefly talks about evolution

comparing human genes with the genes of chimpanzee.

**Least seeked videos (F13) : Total>1430) – Less than 1 standard**

**deviation below the mean ( Arranged in descending order)**

1) Extra4.3-ForensicTechnologies Number of seeks- 0

2) Extra2.7-ModernStrategies Number of seeks- 0

3) Extra2.1-Introduction Number of seeks- 0

4) Extra1.5-TheHumanGenomeProject Number of seeks- 2

5) Extra2.2-HeartDisease Number of seeks- 2

6) Extra1.6-ImprovementsOnHGP Number of seeks- 4

7) Extra1.4-HumanGeneticMapping Number of seeks- 6