

# Exc2A - The Phish from Vermont

*Ido Algom & Natalie Gilboa*

*22/04/2016*

## Introduction

Phish is an American rock and roll band noted for their musical improvisation, extended jams, blending of musical genres, and dedicated fan base. The Band Formed at the University of Vermont in 1983 (with the current line-up solidifying in 1985).

The Mockingbird Foundation is a charitable organization founded by fans of the jam band Phish in 1996 to support music education for children. The website Phish.net is a fan website that formally adopted by The Foundation as a project in 2005. The site had begun in 1991 and served Phish fans for nearly two decades as static HTML pages. It was re-launched as an integrated database - of setlists, song histories, reviews, and more - in fall 2009.

The website contains tables about each song that ever played in show. Every single show the band has been playing during the years has a unique setlist. This information about each song and each show is recorded at the website database for fans.

Let's get some code - preparations

```
require(gtable)
require(ggplot2)
library(gtable)
library(ggplot2)
```

We chose to explore the statistics about the songs. I copied the table about all the songs the band played during the years. The table contains the following columns:

- Song Name
- Original Artist (The band played a lot of cover songs)
- Time Played (At how many shows the song has played)
- Debut (Show date first time played)
- Last Played (Show date last time played)
- Gap (Amount of shows the song hasn't played)

The last show was played in 2016/01/2016

```
yem <- read.csv("phish.csv")
head(yem)
```

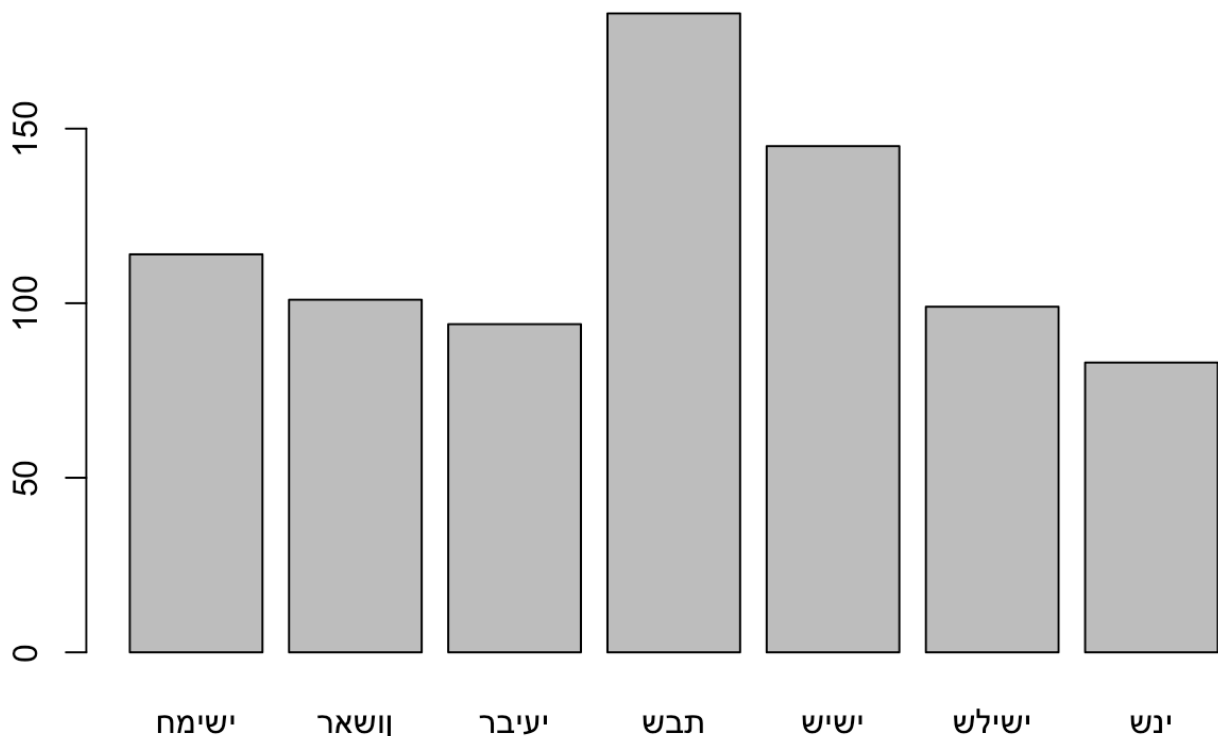
```
##           Song.Name Original.Artist Times.Played      Debut
## 1      Army of One           Phish           17 2003-07-12
## 2      Bathtub Gin           Phish          247 1989-05-26
## 3              Bug          Amfibian           63 1999-06-24
## 4 Crosseyed and Painless Talking Heads           37 1996-10-31
## 5      Down with Disease           Phish          261 1993-12-31
## 6      Farmhouse           Phish           62 1997-11-07
##      Last.Seen Current.Gap
## 1 2016-01-17           0
## 2 2016-01-17           0
## 3 2016-01-17           0
## 4 2016-01-17           0
## 5 2016-01-17           0
## 6 2016-01-17           0
```

## Basic Information

How many songs debuted at each day of the week. In this Graph we took the debut day and converted it to Date and to Weekdays with basic R functions.

The comment line is here to replace the bars arguments as Hebrew in Mac sometimes is not doing such a great job.

```
#barplot(table(weekdays(as.Date(yem$Debut))),names.arg=c("Thu","Sun","Wed","Sat","Fri","Tue","Mon"))
barplot(table(weekdays(as.Date(yem$Debut))))
```



## Songs Left Behind

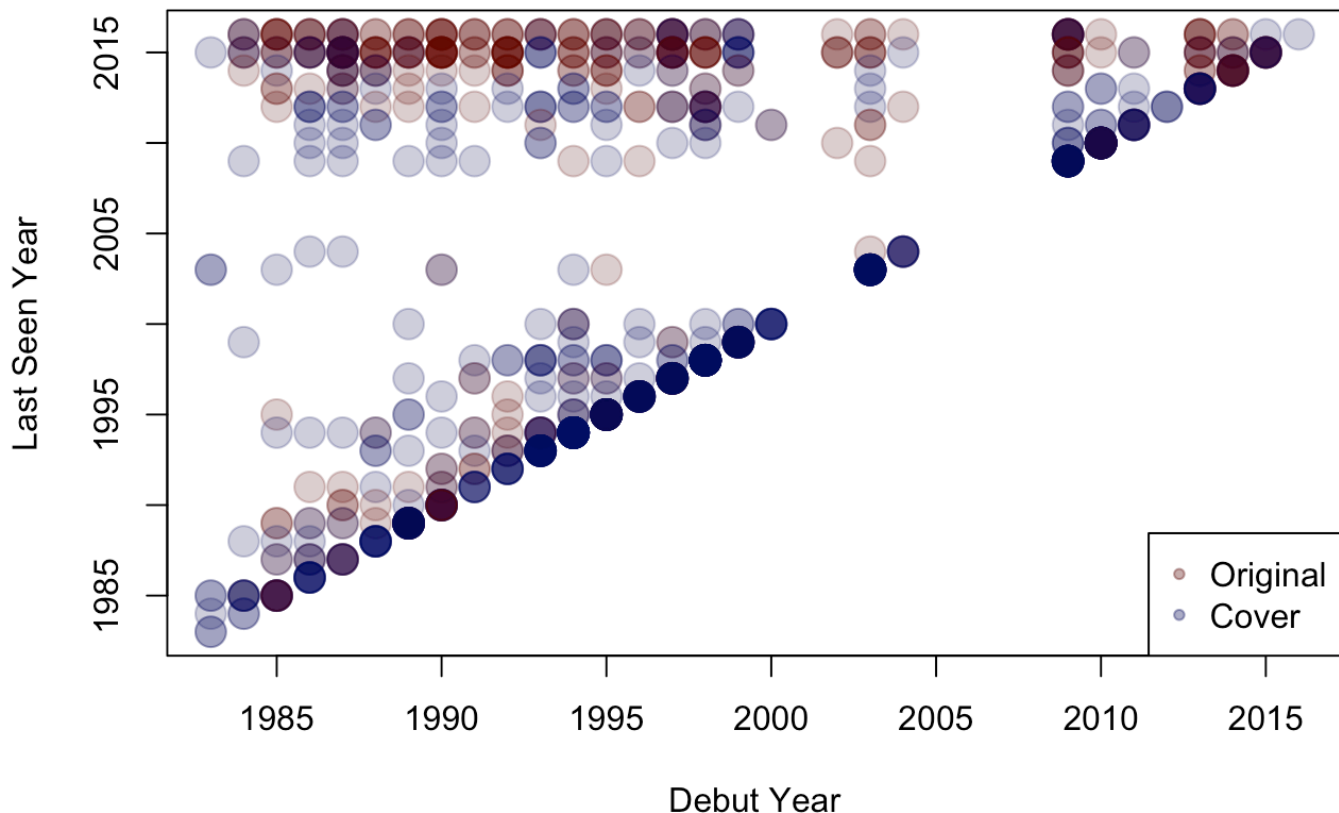
In this Scatterplot we extracted the year from every song, when debuted and last played. The color shows if more songs this year was a cover or original. As stronger the color is, more frequent the group repeats in that point. From this graph we learn that a lot of covers are not repeating themselves. The band play it during it's debut year and then leave it behind. Another thing we can learn is that most of the original songs are not left behind. There're some songs that did left in the past before 1990, but the rest are keeping pooping out at shows.

Some preparations

```
cover <- format(yem$Original.Artist)
cover2 <- rep("Cover",length(cover))
cover2[cover == "Phish"] <- "Original"
```

```
plot(as.numeric(format(as.Date(yem$Debut), "%Y")),as.numeric(format(as.Date(yem$Last.Seen), "%Y")), xlab="Debut Year", ylab="Last Seen Year",main="Songs Left Behind",
pch=20,xlim=c(1983,2016),ylim=c(1983,2016), col=ifelse(cover2=="Original", rgb(100,0,0,60,maxColorValue=255),rgb(0,0,100,60,maxColorValue=255)), cex=3)
legend("bottomright",legend =c("Original","Cover"),col=c(rgb(100,0,0,100,maxColorValue=255),rgb(0,0,100,100,maxColorValue=255)),pch=20)
```

## Songs Left Behind

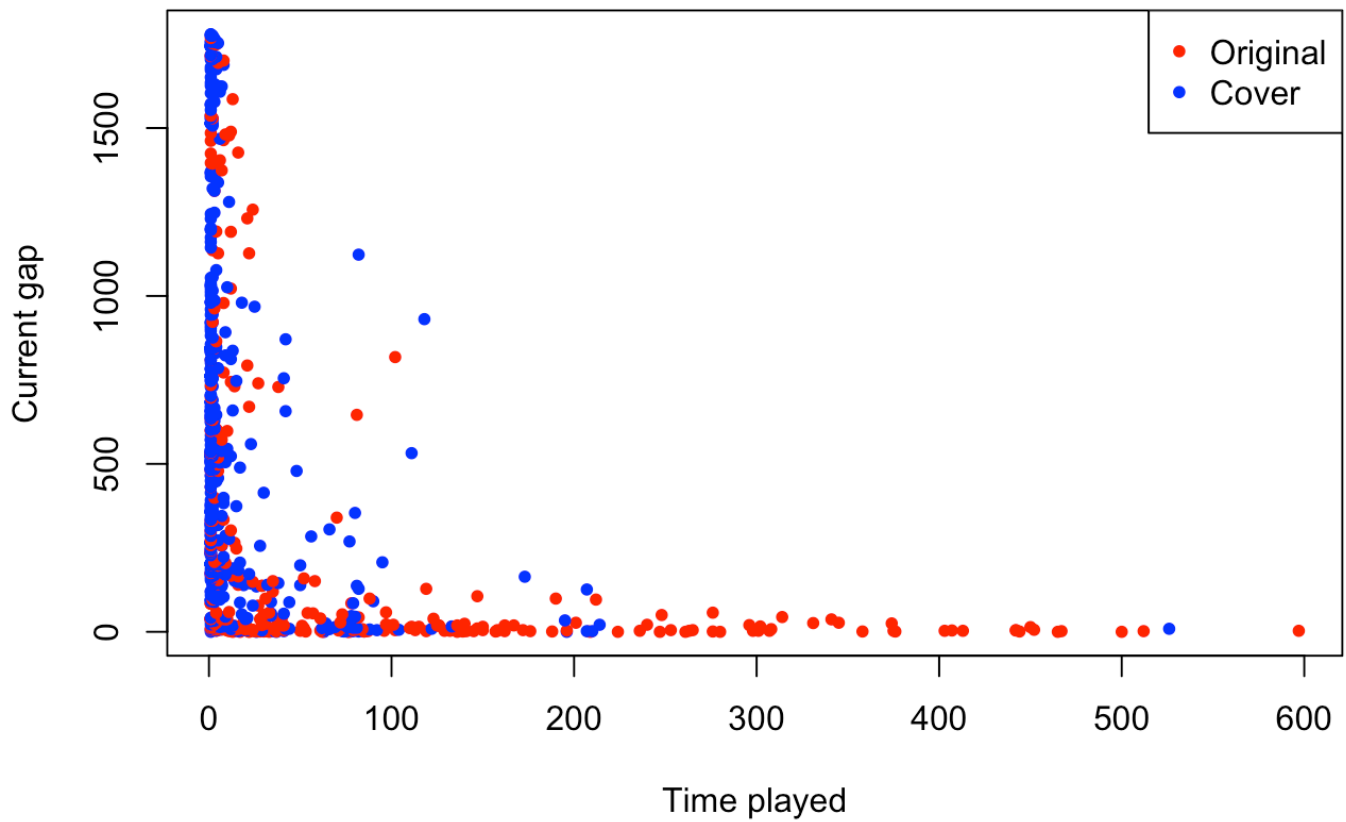


## Bust out

A bust out is a sudden large gap amount of song appearance in a show. From this plot we can learn most of the originals songs are played very often. Most of the covers don't make more than few appearance and there are some songs that used to be played often and now they're gone from the setlist. We can say that as close as a song to  $p(0,0)$  the chances to get played at the coming Summer Tour 2016 are higher. As far a song is on the Gap axis as bigger the bust out will be.

```
plot(yem$Times.Played,yem$Current.Gap,col=ifelse(cover2=="Original","red","blue"),pch=20,cex=1,main="The bust outs",xlab="Time played",ylab="Current gap")
legend("topright",legend =c("Original","Cover"),col=c("red","blue"),pch=20)
```

## The bust outs



Self-Note - Phish is the greatest band for me. A band that changed my entire life and made the person I am. I do