|  |
| --- |
| library(odbc) |
|  |  |
|  | # Create a connection to SQL Server using ODBC |
|  | # NOTE: Be sure to change the database to your actual database |
|  | myconn <- DBI::dbConnect(odbc::odbc(), |
|  | Driver = "SQL Server", |
|  | Server = "ist-s-students.syr.edu", |
|  | Database = "IST659\_M407\_caharper", |
|  | Trusted\_Connection = "True" |
|  | ) |
|  |  |
|  | # Ready the SQL to send to the Server |
|  | sqlSelectStatement <- |
|  | "SELECT |
|  | vc\_VidCast.vc\_VidCastID |
|  | , vc\_VidCast.VidCastTitle |
|  | , DATEPART(dw, StartDateTime) as StartDayOfWeek |
|  | , DATEDIFF(n, StartDateTime, EndDateTime) as ActualDuration |
|  | , ScheduleDurationMinutes |
|  | , vc\_User.vc\_UserID |
|  | , vc\_User.UserName |
|  | FROM vc\_VidCast |
|  | JOIN vc\_User ON vc\_User.vc\_UserID = vc\_VidCast.vc\_UserID |
|  | " |
|  |  |
|  | sqlResult <- dbGetQuery(myconn, sqlSelectStatement) |
|  |  |
|  | # Use +/- 3 sigma to prune outliers (Symmetrically distributed) |
|  | sqlResult <- subset(sqlResult, ActualDuration > 0) |
|  | sigma <- sd(sqlResult$ActualDuration) |
|  | mu <- mean(sqlResult$ActualDuration) |
|  | upper <- mu + (3\*sigma) |
|  | lower <- mu - (3\*sigma) |
|  |  |
|  | # Thanks, Shaun, for this handy outlier fix! |
|  | sqlResult <- subset(sqlResult, ActualDuration < upper) |
|  | sqlResult <- subset(sqlResult, ActualDuration > lower) |
|  |  |
|  | # Create a list of days of the week for charting later |
|  | days <- c("Sun", "Mon", "Tues", "Weds", "Thurs", "Fri", "Sat") |
|  |  |
|  | # Create a histogram of durations (appears in the Plots tab) |
|  | hist(sqlResult$ActualDuration, |
|  | main="How long are the VidCasts?", |
|  | xlab="Minutes", |
|  | ylab="VidCasts", |
|  | border="blue", |
|  | col="grey", |
|  | labels=TRUE |
|  | ) |
|  |  |
|  | # Hard coded fix for hist(): [which(sqlResult$ActualDuration < 400)], |
|  |  |
|  | # Plot a bar chart of video counts by day of the week |
|  | dayCounts <- table(sqlResult$StartDayOfWeek) |
|  |  |
|  | barplot(dayCounts, |
|  | main="VidCasts by Day of Week", |
|  | ylab="Day of Week", |
|  | xlab="Count of VidCasts", |
|  | names.arg = days |
|  | ) |
|  |  |
|  | DBI::dbDisconnect(myconn) |