1. **Intro self**
   1. Who I am
   2. What I do
   3. My background
2. **Very quick intro to R & R Studio**
   1. R – programming language with a focus on statistical analysis
   2. R Studio – IDE to work interactively (or not) in R
   3. Interactive work is where it’s at
   4. Packages – add new functions
3. **Why use R**
   1. Look at the chart
   2. Look at some tasks
      1. Pull complex mailing lists
         1. Use R (or some such)
            1. Save criteria, and let data change underneath
            2. QV has no way to do complex OR logic
      2. Exploratory relational analysis
         1. QV is pretty good for this (until you get into ORs or filtering nulls)
      3. Statistical analysis
         1. Use R
            1. QV just doesn’t have the tools
      4. Creating a report for others
         1. Use R
            1. Save criteria
            2. Write with R Markdown
         2. Or QV
            1. Unreliability of filters
      5. Predictive modeling
         1. Use R
            1. QV has no tools for it
4. **Tidyverse**
   1. A meta package designed to simplify working in R
      1. Includes a bunch of packages including dplyr, tidyr and ggplot2
   2. The pipe
      1. Little bunny foo foo
   3. Building lots of little objects
      1. Like re-usable SQL subqueries
      2. Don’t worry about memory usage
   4. Dplyr verbs
      1. mutate() adds new variables that are functions of existing variables
      2. select() picks columns based on their names.
      3. filter() picks rows based on their values.
      4. arrange() changes the ordering of the rows.
      5. group\_by() allows you to “pick pivot table rows”
      6. summarize() reduces multiple values down to a single summary.
   5. Database connections and relational verbs (see slide 6 & 7)
   6. Fancy plotting
      1. Dates vs. Arrival delay
      2. Dates vs. Arrival delay w/ y limit & jitter
      3. Dates vs. Arrival delay faceted by carrier
5. **Summary**
6. **If we have time – look at my code**