By James Schacht - Westminster Heights Resident

The woman buried under this headstone was born in the same century as the civil war and died in the same century as the iPhone. Imagine the changes she saw in her lifetime! From societal and cultural to scientific and technical, she experienced dramatic changes in the human condition. She is interred, along with many other area residents, in the “cemetery at the end of the street”.



## Preamble: Respect for People Living and Dead

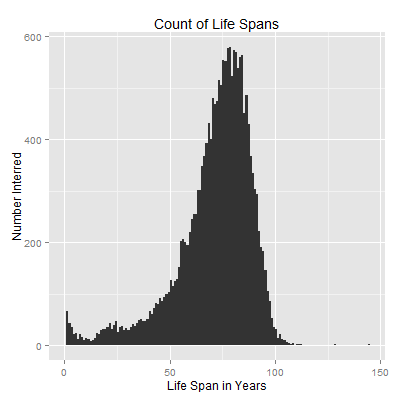
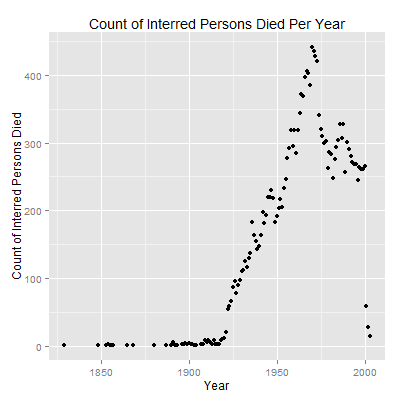
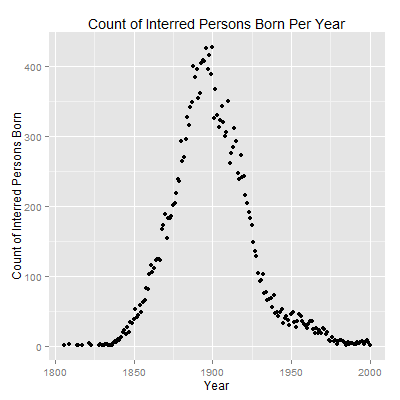
Data analysts sometimes appear to be cold when referring to information, knowledge, and/or wisdom derived from data. This report presents information based on a data set comprised of data about real people who have lived and have relatives and descendants who are still living. The author understands this and intends only the highest respect to each of those individuals and families. The information presented considers the whole of the data mostly; although outliers become evident, such as oldest or youngest person. The goal is to present the facts about the data, with respect to the persons, in order learn from it.

## Data Scientist Intrigued by Data that Touches Daily Life (and Death)

I became fascinated with this cemetery and the people interred there while walking my dog there daily. A quick search on the internet turned up a list of interments generated by a canvassing effort led by the Pinellas Genealogical Society and executed by an Eagle Scout earning a badge in 2000 and 2001.

To a data scientist this list is what we refer to as “dirty data”. Called that because in its raw format it is hardly useful beyond a simple text search due to the way the data was captured and presented. There are many problems due to inconsistency in the information captured, the columns in which it was recorded, and human error. Thanks to excellent tools, I was able to programmatically “clean-up” the data and provide insight into the information and knowledge hidden there. More work is needed in this area, but we can already use this insight for education, historical research, planning and predicting.

## Quick Insights

* There are 22,079 records in the list. Logically we have to assume there are fewer interred than this number due to some folks not being buried in plots they purchased (re-marriage, etc.)
  + Interment density = 401 per Acre
* There are 7,533 *unique* surnames in the list. While not everyone with the same surname is related, it suggests that there is an average of about three persons per family interred here.
* There are two people registered as 144 years old at death. This seems incredible, but that is what was recorded.
* There are 94 persons more than 100 years old interred.
* Average life span of those interred 70.76 years.
* There are 108 persons less than 2 years old interred.

## Work in Progress

I met with the Pinellas Genealogical Society (PGS) <http://www.flpgs.org/> and learned that they have similar data for every cemetery in the county. We are looking into a project to enable the systematic cleansing and presentation of this data in a format that will be accessible to genealogists, historians, educators, and the public.