Introduction

Recent years have seen many changes in family structure in Western countries. Lower birth rate, increased number of single parent and same sex families, the statutory recognition of unmarried cohabitants (e.g., common-law marriages), and the growing number of non-religious weddings are all indicators of change in family values in Western society.

While different explanations have been proposed in order to elucidate what the impetus behind this change could be, the focus of our work is to try to see if these changes are reflected in publicly available data collected by the US Census Bureau.

Data and Process

The data we used for this work is part of the “Families and Living Arrangements” dataset available from the US Census Bureau [1]. This data is based on a yearly population survey collected every March. The dataset explores household composition from different angles such as whether both parents are present, whether the children live with their parents, what is the family income level, what is the race and ethnicity of the parents, etc. One drawback of the data is that it is limited to recent years, during which the data collection granularity has varied. For example, in 2005, data was collected separately for the 45-49 and 50-54 age groups whereas in 2002, data is only available in aggregate for the 45-54 age group. Due to this and to our emphasis on trends that are exhibited by the general population (non-race related), we concentrated on data for the years 1999-2005 and restricted ourselves to data that was available for all these years. In other words, we focus on the intersection of age group ranges of different years, after we had coalesced certain age ranges.

Our dataset provided separate counts for males and females 15 years and older dividing them into one of the following marital statuses:

- Married Spouse Present
- Married Spouse Absent
- Widowed
- Divorced
- Separated
- Never Married

As an example of a person belonging to the “Married Spouse Absent” category consider a person married to a deployed Armed Forces member. For our purposes, the distinction between “Separated” and “Divorced” is not of great importance.

The dataset that was available online was in Comma Separated Value (CSV) format, each year represented by a different file. We took this raw data and loaded it into Microsoft Excel and collected only the data of interest. As explained earlier we had to redistribute the data into new age group ranges,
and this was done semi-automatically using cell formulas. Naturally, we had to recalculate the percentages for each age group from the redistributed absolute data.

Since our “Families and Living Arrangements” dataset provided us only with aggregate data for 7 consecutive years, we felt that in order to do true justice to the TimeSearcher [4] tool, we have to also consider data that fluctuates much more and spans over many more time points. For this purpose, we considered an (unrelated) dataset published by the US Department of Energy containing the Daily Spot Prices for various Petroleum products [2]. This dataset contains the daily spot prices from 1986 to this day at different important stations along the oil supply chain. In order to filter the data we imported it into Microsoft Access and ran several SQL queries on it. The result was exported to CSV format and then imported to TimeSearcher.

Findings

✓ Surprising number of men lived away from their spouses in 2001
✓ Divorced men tend to remarry more than women!
✓ Remote natural disasters have a stronger impact on oil prices than major local events
Surprising number of men lived away from their spouses in 2001

Figure 1: In 2001, sudden increase in the number of males 20-64 living away from their spouses

Figure 1 shows a spike in the percent of males aged between 20 and 64 who live away from their spouses. Although this tendency does not appear to be coincidental, because we had no reasonable explanation for it, we went back and rechecked our original dataset and we can indeed confirm that this pattern is also present there.
Divorced men tend to remarry more than women!

Figure 2: Men 20+ who never married (in blue)
Figure 3: Women 20+ who never married (in blue)

Figures 2 and 3 show that for the same age groups, more men than women were never married. In this consideration we excluded people in the 15-19 age group in which marriage is significantly less common.

Figure 4: Men 20+ who are divorced (in blue)
Figure 5: Women 20+ who are divorced (in blue)

Conversely, for the same age groups, more women than men are divorced. This result suggests that divorced women tend to remarry less than do divorced men. One possible explanation is that the number of men who wish to get married is smaller than the number of such women. Another possibility is that women may be avoiding remarriage because of psychological reasons. Other reasons could be that these women may have already fulfilled their expectations from marriage, or they are just faced with a smaller pool of eligible men.
Remote natural disasters have a stronger impact on oil prices than major local events

Figure 6: Oil Prices around the time of Indian Ocean Earthquake and the following Tsunami
Figure 7: Oil Prices around the time of Hurricane Katrina

From figures 6 and 7, it appears that natural disasters can have strong short-term impact on oil prices. We also investigated the possible influence of human-induced disasters such as the 9/11 attacks, the Space Shuttle Columbia crash, and the Northeast power blackout of 2003. However, in these cases, we observed no such correlation.
Another event that was suspected of having an influence on oil prices was the Midterm Elections in 2006. The media raised questions as to whether it could be that oil prices were deliberately lowered in order to bolster the ruling Republican party [5]. While we observed some decrease in oil prices in the days before the elections, our study neither shows conclusive evidence to support this claim, nor to refute it. In this context, it is important to note that we did not have data about gas prices to the end consumer.

**Tool Critique**

TimeSearcher 2 proved to be a useful tool to visualize the change in data over time. But we believe it can become even more effective with some enhancements.

TimeSearcher 2 allows users to see a graph with all of the items in the dataset. The users can filter some of the items that they do not need. TimeSearcher 2 can also display a set of graph plots, where each graph plot corresponds to one variable chosen by the user. These plots are tiled vertically essentially sharing the same timeline.

Being a visual exploration and analysis support tool, TimeSearcher 2 empowers the user with primary control over the parameters of the visualization. Thus, the initiator of every action is the user. TimeSearcher 2 caters also to the needs of more advanced users through the availability of shortcut keys. Users can avoid the time spent on accessing menu items by using the key combination that corresponds to the desired task.

In our opinion, the biggest obstacle to using the tool is formatting the data correctly. Since users gather data from disparate sources, converting it to the proprietary \texttt{tqd} file format can be challenging.
Moreover, TimeSearcher only partially supports the \texttt{tqd} format. For example, it does not allow the use of two static variables, although this is supported by the \texttt{tqd} files. One feature that makes part of the conversion easier is the good support TimeSearcher 2 has for handling missing data points.

While TimeSearcher 2 does allow zooming to a specific time interval, i.e. zoom on the horizontal axis, there is no corresponding way to zoom vertically. When the data points lie very close to each other, the user may find it hard to distinguish the two curves. Users are also unable to plot two or more variables on a single graph. This can make comparison between two variables to be unnecessarily difficult to perform.

TimeSearcher still does not provide context-sensitive help. Users that need help have to read through the Quick Reference or the User’s Manual. As a user, it would be helpful to be able to access only the relevant portion of the documentation by just pushing the F1 key. Also, users only have limited opportunity to reverse their actions as the only undo action is to reverse the filtering of the unselected items. Furthermore, that undo button, “Undo Filter Unselected Actions”, is always enabled even if there is nothing to undo. This can be pretty confusing.

TimeSearcher 2 allows users to choose to view only the selected items. However, as users go back and forth between viewing only selected items and viewing all items, TimeSearcher 2 always starts by showing the graph for the first variable independent from the current variable selection. In a way, TimeSearcher 2 ‘forgets’ what is selected while switching to another view.

As a final observation, it is worthwhile noting that TimeSearcher 2 crashes frequently due to uncaught exceptions, leaving the user with a cryptic error message. Besides catching runtime exceptions in order to avoid crashing the whole system, we believe that a relatively lightweight software such as TimeSearcher can afford using an error logging and reporting mechanism without suffering a critical performance penalty.

In this project, we saw no clear advantage to using TimeSearcher 1 over TimeSearcher 2, for the type of questions we wanted to investigate. Generally, we felt that the unique features of TimeSearcher 1, namely time queries, were not necessary for our research questions.

References


