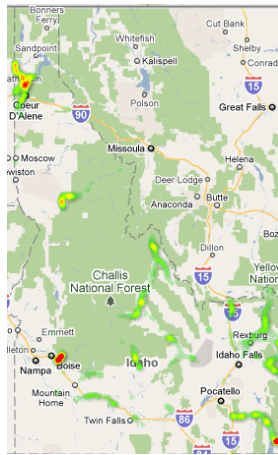# Idaho Highway Wildlife Mortality

A. James Frankman

Abstract—Idaho wildlife mortalities on highways and roads is tracked by the Idaho Fish and Game and the data is made available to the general public through an API called IFWIS Core. While the data supplied does offer species information and geographic coordinates, it can be difficult to organize and understand. This paper will attempt to organize and present this data in visual form using Google Maps and Visualizations APIs to show facets of wildlife mortality in Idaho by density of occurance, time of year, and species variety

Index Terms—Information Visualization, Idaho Fish and Game, IFWIS Core, road kill, wildlife mortality, Google Visualization API.

✦

1 INTRODUCTION

Amongst the rural communities throughout the United States, the attrition of wildlife by highway collision is a common occurrence. In an effort to better track and understand wildlife collisions occurrences, the Idaho Fish and Game tracks highway collisions that have occurred since 2001. This data can be useful and relevant to several areas of study.

First, understanding how and where collisions occur can help prevent traffic accidents. According to the National Highway Traffic Administration 4% of all traffic accidents in the United States are collisions with wildlife[1]. The collisions with wildlife on U.S. roads and highways represent a significant safety concern to motorists. Besides the risks posed to motorists, the affect on wildlife populations is also significant. America's wildlife is a natural resource, and highway collisions have a negative impact on wildlife populations. In addition, the time and place of wildlife collisions can reveal trends such as common migration routes and wintering areas as well as the diversity of wildlife itself. For instance, one can infer the health or population of from areas with a high occurrence of deer collisions. It can also reveal intrusions of species into areas previously uninhabited, for, if a Gray Wolf is killed on a freeway in an area that previously had been devoid of Gray Wolves, one can easily conclude that Gray Wolves have entered the area.

The Idaho Fish and Game offers the public access to wildlife deaths on roads and freeways in two primary ways. First it makes the data available via API for software developers and web authors. Second, it offers a Dashboard application to view the latest 250 incidents. However, this data has several limitations that could be overcome with certain information visualizations.

One problem with the data is its overall breadth. There are currently over 11000 records of wildlife collisions and when all of these are displayed on a map the information overwhelms the viewer. For instance even from the Idaho Fish and Game dashboard, which

only shows 250 of the latest observations, the density of markers on the map make it difficult to distinguish individual incidents.
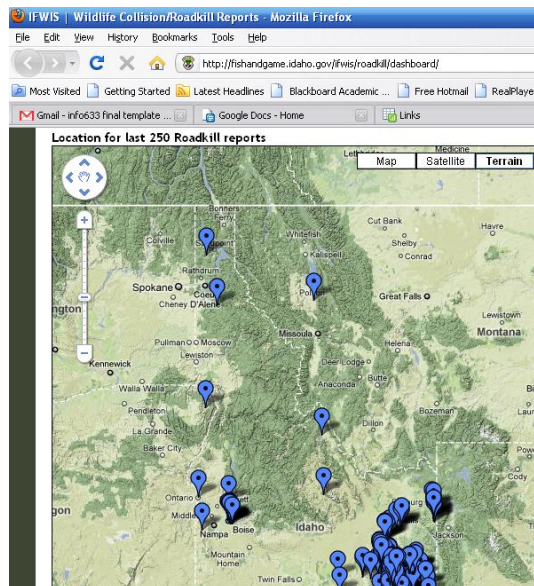
There is no way for a viewer to easily zero in on road kills for a particular species or time of year. The data markers are not distinguished in any way to show the type of incident according to species or time. All makers appear identical and the viewer must click each one to find out the nature of the incident. In addition, the

markers on the map are so large it only takes 8 or so to saturate the map canvas. Thus, an area containing about 8 incidents will appear the same as one with dozens more.

A final problem is the accessibility of the data to the public at large. While the data is available via API, it takes persons knowledgeable with web programming to write their own web pages and construct their own queries to filter the reported incidents to obtain the view they are interested in. One viewer may be interested in high occurrence areas while others may want to satisfy curiosity and view uncommon species or rare occurrences.

2   SOLUTION AND APPROACH

The raw data maintained by the Idaho Fish and Game and made accessible by the Idaho Fish and Game's IFWIS Core API and website can be effectively organized for in-depth analysis using the Google Visualization API in combination with Google Fusion Tables services and API's.

2.1   Google Fusion Tables

Google Fusion Tables is an online service where users can upload, share, and merge data. Data is organized into tabular form with tables and views similar to how data is organized in a traditional database. However, it can be done so without in-depth technical knowledge by the user.

Google Fusion Tables require a Google Account and a basic understanding of structured data. Upon logging on with their Google account, users can upload their own data to create tables or can browse and find publicly available tables shared by other Google Fusion Tables users. Data can be uploaded in various forms including comma delimited files (.csv), xml, kml, Microsoft Excel spreadsheets, and Open Document Standard spreadsheets.

Once data is uploaded, Google Fusion Tables will create a table based upon the data uploaded and will organize the data into rows and columns. Each column within a table can be classified by type including text, number, Location, etc, as well as format, such as a link or image. Given the data for Idaho Wildlife collisions was

location based, the included column type of location was particularly useful since it allowed the data in the table to be rendered on a map almost effortlessly.

The Google Fusion Tables offers several ways to organize data such as filtering data by certain query parameters and aggregating data by column. Probably one of the most useful features of Google Fusion Tables is its merge capability. With this feature data contained into two separate tables can be combined by a
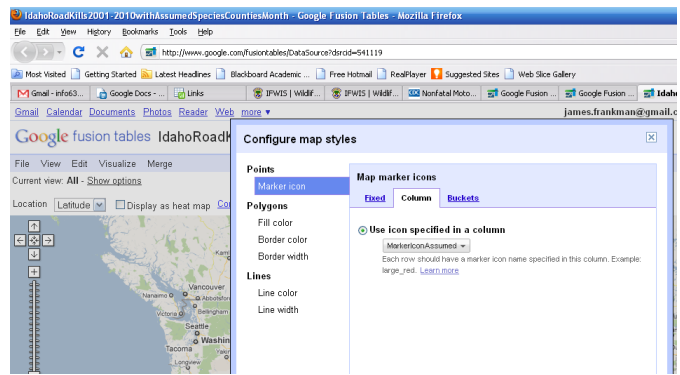
common key. For example, data found in an existing fusion table containing the names and geometric shapes of US counties could be combined with the wildlife collision data according to County Name.

Google Fusion Tables was a primary means of overcoming the current limitations with IFWIS Core data as it allowed the data to be accessed and organized more easily than what was possible through API service calls. Some of the techniques used to accomplish this are explained in section 3 of this document, Data.

2.2   Google Visualization API

The Google Visualization API is a tool whereby data can be rendered visually using charts, graphs, maps, timelines, trees, and other visual components. For the purposes of the research covered by this paper, the API's map components were very important [3].

At its core, the API does require Javascript and web programming. However, many of the API's features were seamlessly integrated with Google Fusion Tables. This removed the need for manual programming to produce many of the visualizations found in this paper. Instead of manual Javascript programming, the API's features could be accessed from the Google Fusion Table's *Visualization* menu. From this menu one is able to produce various geographic, chart, motion, and timeline visualizations.

Once a particular visualization is chosen, the user is given even more access to the Visualization API's features through menus and other web controls. Because of this, non-programmers are able to take advantage of the Google Visualization API without programming experience by using the built in features within the Google Fusion Tables API.

3   DATA

As mentioned previously, Idaho wildlife collision data was obtained from the IFWSI Core API service. The data can be downloaded in various forms and types. Since the data needed to be uploaded to Google Fusion tables, data in comma delimited form was the best choice since the csv format takes the least amount of space. The data was downloaded using the IFWIS Core API using a typical http get operation. In the URL the desired format of the data could be specified along with query parameters to direct which records should be returned. Below is an example of the URL used to retrieve records for 2001 wildlife collisions:

http://fishandgame.idaho.gov/ifwis/core/view/roadkills/2001.csv?species=0&start=&end=&highway=0&mpFrom=&mpTo=&nstart=01/01/2001&nend=12/31/2001&pageSize=2000

3.1   Problems within IFWIS data

There were several problems with the IFWIS data in terms of content as well as accessibility. A short description of each of the problems is found below.

3.1.1 Accessing Bulk Amounts

The first obstacle came from obtaining IFWIS data came from accessing the data itself. Since the data was exposed via API it is likely that the intended purpose of querying and accessing the data was for small interactive sessions rather than bulk data exports. When specifying a query to download a large dataset, timeouts and other errors occur. As a result, the data needed to be downloaded in smaller chunks. To do this, 11 queries were run, one for each year from 2001 to 2011. These queries collectively downloaded into 11 separate comma delimited files, one for each year. Once all 11 files were downloaded, they were combined into a single comma delimited file using simple text editing tools (Notepad++).

3.1.2 Inconsistent Data

Once data was combined into a single comma delimited file, it was uploaded into a Google Fusion Table. However, it became clear that much of the data was incomplete, inaccurate, or dubious. When plotted on a map, some of the records were shown outside of Idaho Boundaries. This indicated a mistake in the record's longitude or latitude data. Some records also had incomplete or erroneous data. For instance, there were several hundred records of collisions where the species name was incomplete and contained the genus for deer: *Odocoileus* but did not contain the species designation *hemionus* or *virginianus* which would indicate either a Mule Deer or White Tailed Deer. Besides problems with the species names, there were a few records which appeared dubious. For instance, there was one record for species *Dasymutilla* which is a genus for ant. It is likely that a member of the public recorded the incident on the IFWIS website as a joke. There were also several hundred records which had the species as "unknown".

  In order to fix the problems with inconsistencies, assumptions were made on incomplete species names. For instance, in the case of only the genus of deer being reported, these records were assumed to be Mule Deer since these are the more common species in Idaho in most areas. The data was loaded into a spreadsheet and a formula was used to modify incomplete species names to their assumed name and put the assumed name in a new column. Dubious records such as those containing "Ant" or "unknown species were tossed out and not included in the fusion table.

3.1.3 Only Longitude and Latitude

Many records that did contain species and other data were missing longitude and latitude data at all, making impossible to even plot these on a map. Although many of these did contain species information, they were not included in any of the geographic visualizations. The primary focus of the research was to show geographic occurrences of wildlife collisions, therefore any records that could not be plotted on a map were not included. However, they were included in the calculation that counted number of occurrences by species.

3.1.4 Common name missing

The data downloaded for IFWIS website only contained the scientific names for species. This was problematic because a record being displayed on a map would only show the user *Odocoileus hemionus* rather than the commonly understood name of *Mule Deer.*

  In order to address this problem, features in Google Fusion Tables were used to include the species common names in the fusion tables. To do this, the table was aggregated by species name, this operation returned a result that only showed a single record for every distinct species found in the table. Next, a new table was created based upon these aggregated results having the columns of scientific name and common name. The common name for each of the entries in the table where then entered. Once this table was complete, it represented a map between scientific names and species common names. This table was then merged back into the full table containing all wildlife collisions with the resulting table containing all columns found in the original table as well as a new column for common name.

3.1.5 Lack of Pictures

One problem with both the data and the visualizations supplied on the Idaho Fish and Game dashboard was there were no pictures for each of the species recorded. While this may not pose a problem to wildlife experts and sportsmen, to the general public who may not have seen the species know what they look like, this is a significant problem.

  Two columns were added to the wildlife table to hold links to photos. Photos for each of the species were obtained from the CAL University photo database[4]. In this case, the data was again downloaded to a spreadsheet and a formula used to set the image link value to point to the proper photo of the species contained on each record. This could have also been accomplished with a fusion table merge, but using a spreadsheet was chosen for this operation instead.

3.1.6 Lack of County Information

 All records contained in the IFWIS database only included longitude and latitude for location information. This is important for being able to plot the records against a map, however, the lack of any other location information is problematic when producing visualizations for geographic locations such as counties. The Idaho Fish and Game provides downloads for its own game management units that is make available in kml format. This kml file contains the geographic

shapes that can be plotted shown as a layer on a map. There was also kml information for United States counties. However, joining the collision data with counties or game management units, was impossible since the IFWIS database did not contain any fields indicating county or game management unit. The problem with the data proved most problematic for the entire project as a primary goal was to create a visualization that could show kill density per geographic area.

In order to inject county information into the IFWIS data, a process had to be created whereby latitude and longitude coordinates could be used to determine county information or Idaho Game Management information. Unfortunately, a service or data source could not be found that would provide cross referencing between geographic coordinates and IDFG game management units. However, counties were able to be discovered using the Google Maps API to find each record's county according to geographic coordinates. Within the results returned by the Google Maps API reverse geocoding service, administrative units such as States and counties are returned. The service is freely available and quite easy to use when reverse geocoding small bits of data. But, the IFWIS data records contained thousands of records and as a result, considerable time had to be devoted to create a script that could reverse geo-code thousands of records at a time. The production of this program took the bulk of the programming needed for this project. It would have been more ideal if a method could have been used to join longitude and latitude coordinates to geometric shapes contained in the KML files without the need for an intermediate reverse geocoding step. However, a method for doing so could not be found.

Once all records were reverse geocoded, results were recorded into a comma delimited text file that contained the row key for each IFWIS record and county. This file was then uploaded to Google Fusion tables an merged with the IFWIS fusion table.

At this point the data needed for the visualizations was in a proper state to allow for interactive visualizations that could show species photos and common names as well as show visualizations according to county.

4 VISUALIZATIONS

### 4.1 Wildlife Collisions Interactive Visualization

One of the problems with the Idaho Fish and Game dashboard as well as the data itself is that it lacked any means of distinguishing between types of collisions when they are displayed on map requiring the user to randomly click each map marker in order to determine details concerning the collision incident. Even once the incident is clicked, no visual photo is displayed to give casual users an idea of what the species looks like. Google Fusion Tables combined with built-in Google Visualization API features was used as a means of showing distinguishing characteristics of collision incidents on the map as well photo information when an incident is selected.

The Google Visualization API allowed marker styles and colors to be specified according to values in the Google Fusion table. Five colored markers were chosen to indicate species type. A specific marker was designated for White Tailed Deer, Mule Deer, Moose, and Elk as these were both fairly common and result in significant

damage and personal injury to drivers. All other species were given a basic yellow marker.
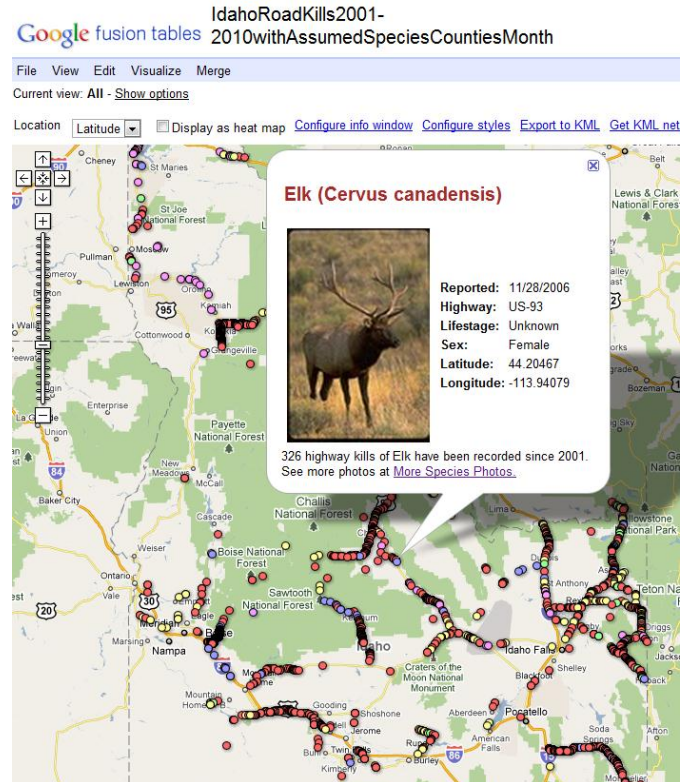
This technique helped highlight the most dangerous and common sorts of collisions while still showing other collision types for the end user. When a maker is clicked data from the fusion table was rendered using html mark-up to show common name, species, photo, and other information. This visualization not only would be effective for casual observers interested in the different types of collisions in
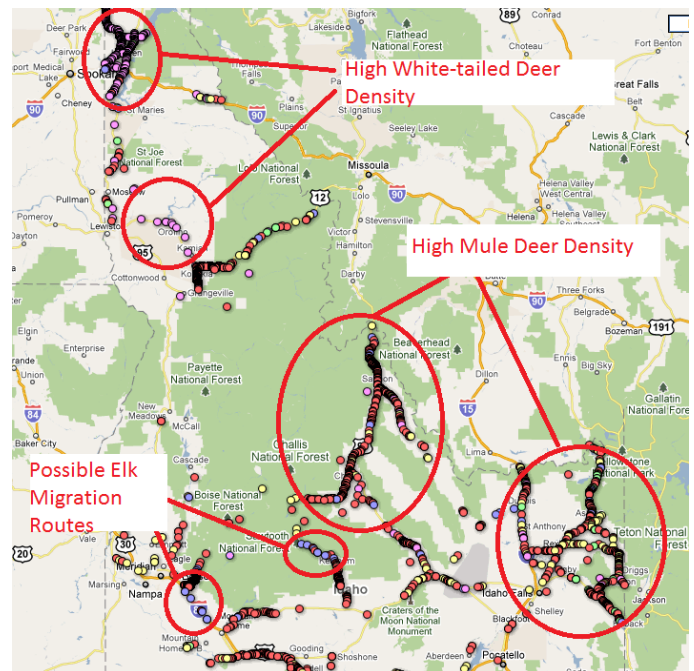

Idaho and where they happen, but also could be used to identify species density or crossing routes. For instance, although only 300+ collisions with Elk were recorded, multiple occurrences of collisions could indicate a denser population or possible migration routes.

4.2 Collisions by Month

When combined with Google Fusion Table's filter and query capabilities other visualizations could be created to answer not just where collisions occur most often, but when. Several different filters were applied to the base data according to month of year. For example, views of the data containing all collisions occurring in July were created and then combined with the visualization and then compared to similar views of collisions in November. The results were markedly different. Times of the year when most big game animals have already migrated to higher elevations such as July have a much lower incident of collisions than during late fall and winter months such as November where big game is forced to move to lower elevations due to heavy snowfall.

The visualization did show that collisions generally occurred in lower elevations. However, there were some contradictions to this such as the in the Island Park/Yellowstone areas of the State.
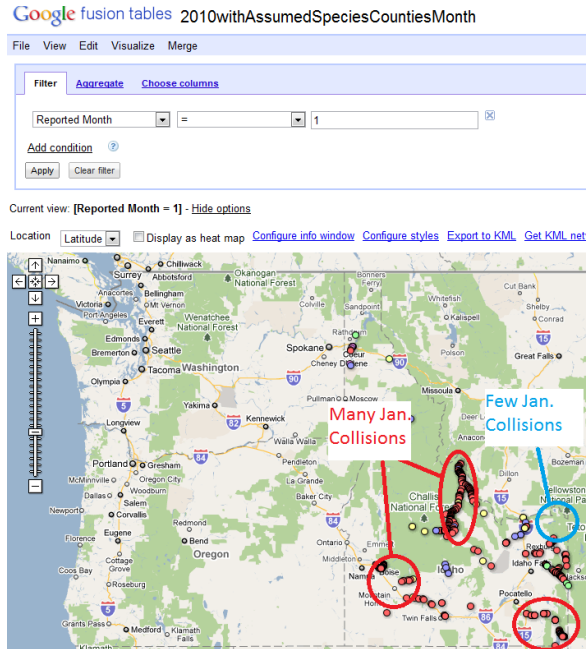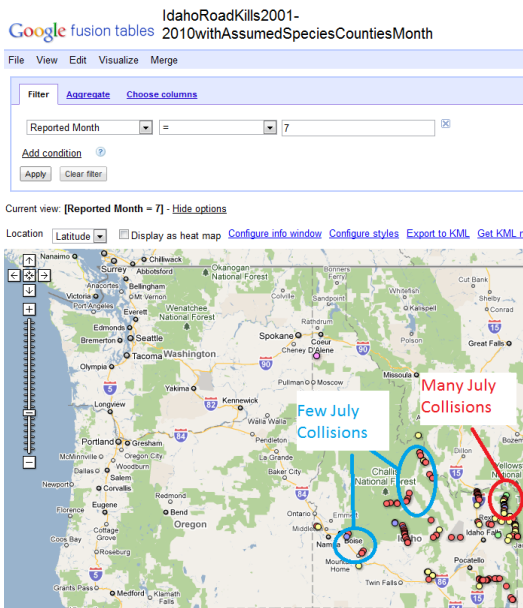
4.3 Collisions by County

Another visualization of interest is Wildlife collisions by Idaho County. The problem with previous visualizations is that there were no boundaries other than the state of Idaho to consider. Collision data could better be represented when organized by county. This was made possible due to the reverse geo coding process described in the data section. Again, Google Fusion Tables combined with the Google Visualization API proved useful for producing a visualization to represent number of incidents per county.
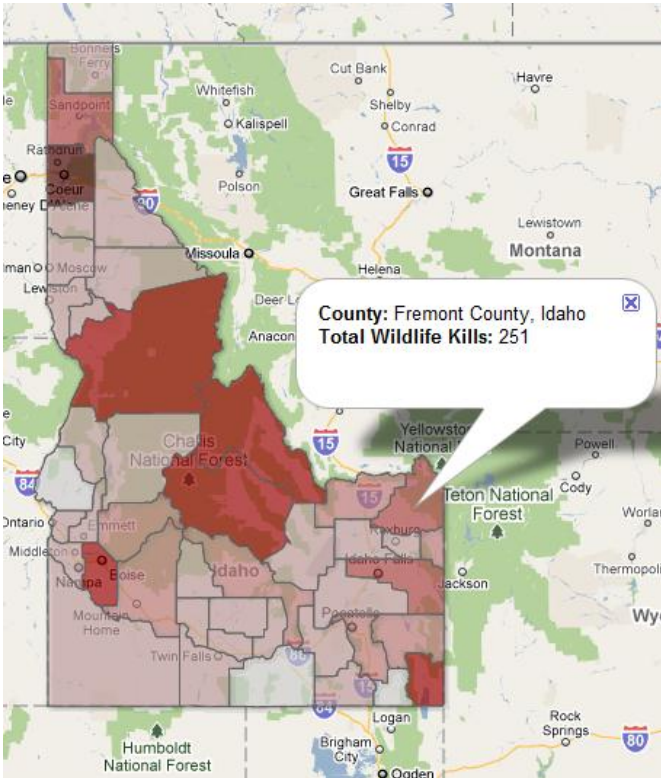
The visualization was created using the Configure styles link. First a new fusion table was creating by aggregating all records by county name. This produced a simple table with county name and number or records for each county. Next this table was merged with the United
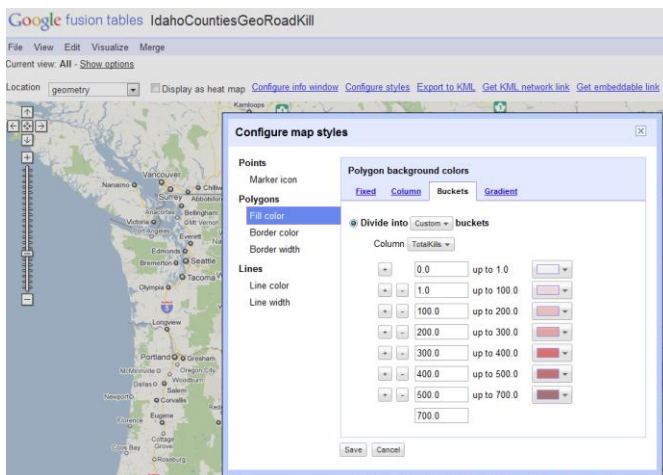
States counties fusion table containing the geometric shapes needed. Once the table was merged it contained both the geometric shapes of the counties as well as the number of collisions for each county. When visualized on a map, the color of each county could be specified according to the total number of collisions per county.

Kooteni, Custer, Lemhi, Idaho, Ada, and Bear Lake counties were shown to have the highest incidents of wildlife collision. However, most other counties in the state show at least 100 collision incidents.

4.4 Other Species

Another perspective of wildlife collisions can be taken in terms of species diversity. While the majority of incidents recorded are collisions with Deer, Moose, Elk, and Antelope, there are records containing less common species. A visualization was created by filtering out certain species to produce a map show all other species types. From an interactive perspective, it can prove interesting to see the various species of wildlife throughout Idaho.


Other collision data may also reveal trends previously unknown. For instance biologists and other experts commonly assert that Gray Wolves keep to the wilderness areas and avoid human contact. However, since 2001 there have been 11 recorded collisions with Gray Wolves. This is far higher rate of occurrence than other supposedly more populous predators such as Black Bear, Grizzly Bear, and Mountain Lion. This may indicate that either wolves have moved into populated areas (contrary to expert opinions) and/or the population of wolves has grown much than is currently being reported.

4.5 Collision Density

Displaying incident density in small geographic areas is one of the biggest shortcomings of the dashboard site. The problem is that even if smaller makers are used, they only can effectively show a dozen or so incidents in a small area. Some areas have hundreds of incidents, yet it is impossible to distinguish these from other areas with much sparser collision rates. Initial investigation into the Google API revealed that one possible approach would be to use expandable markers[5]. However, the problem with this solution is that it would require using a proximity algorithm to calculate the density of markers within a given area. Furthermore, this particular visualization assigned whole numbers as the

measure of marker size. The approach was abandoned in favor of one which could show more granularity.

Fortunately, the Google Fusion Tables exposed a Google Visualization API feature called *heat maps*. To display a heat map all that was required was to open a map and select a checkbox to render the markers on the map as a "heat map"
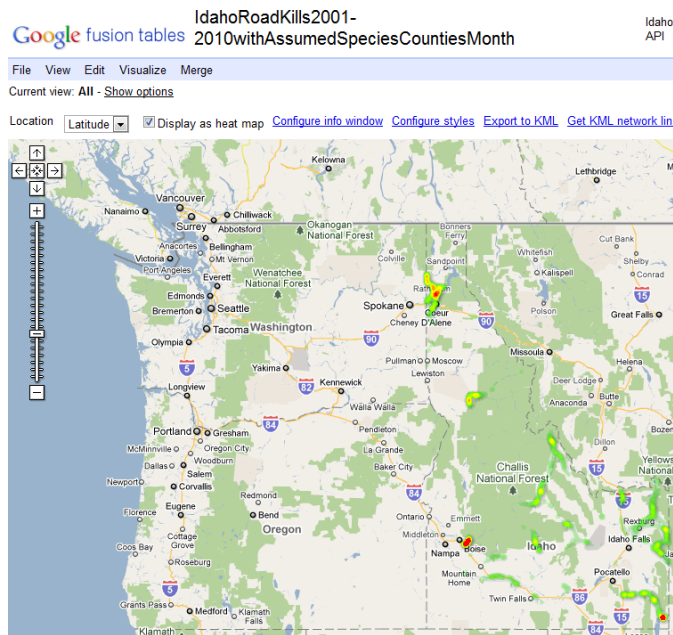
The *heat map* visualizationrevealed several areas of high density collisions that were unclear when viewing all the incidents plotted on a map. For instance, the area around the city of Boise did not stand out from others within the State. However, the heat map visualization revealed that it had one of the highest occurrences of collisions.

When combined with a topographical map, the hot spots in the heat map also reveal the geographic features where high density collisions are likely to occur. Most collision hot spots appear in areas that transition from mountains to lowlands. These areas would indicate zones with the highest safety concerns as well as the zones which have the highest ratios of human/ wildlife contact.
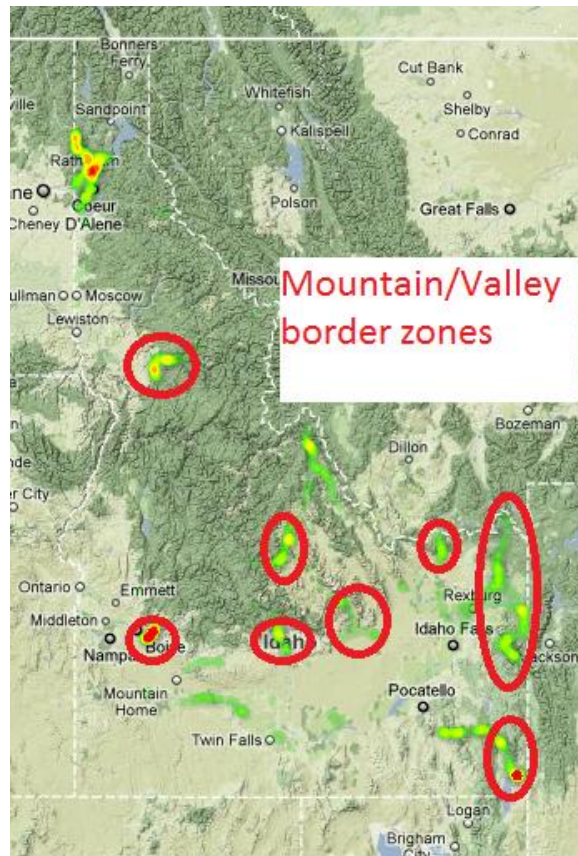
The heat map visualization also reveals collision density according to season of year when combined with the Google Fusion Tables filtering capabilities.

Winter: December, January, and February
Winter generally contains the highest number of collision incidents, particularly around Boise, and in the southeastern corner of the State southeast of Pocatello. However, Yellowstone area north of Rexburg has a low number of collisions during the winter months.

Spring – March, April, and May

Wildlife collisions drop off in the spring throughout the State. The may indicate that wildlife have begun to move to higher elevations and away from roads as snows melt. Or, it may indicate that the density of wildlife have decreased through the normal winter mortality causes such as starvation and disease.

Summer – June, July, and August

Generally speaking, the summer months contain the least amount of highway collisions. The one exception to this is that Summer has the highest rate of collision in the Yellowstone area (northeast of Rexburg). This can be attributed to the fact that the Yellowstone area is devoid of deer, elk, and antelope during winter and early spring due to heavy snowfall. Another contributing factor to the higher number of collision incidents may be due to the increase in traffic around the Yellowstone area due to the summer tourist season.

Fall – September, October, and November

Collision rates begin to rise again throughout the State during the fall. In particular, the areas around Yellowstone either hold steady or increase during the Fall. This is likely due to deer, elk, and antelope migrating from the Yellowstone area to lower elevations as snows accumulate and winter draws near.

5 CONCLUSION

The Google Fusion Tables combined with Google Visualization APIs both proved useful in analyzing wildlife mortality on Idaho's roads and freeways. Google fusion table provided several tools that helped streamline the data importation and scrubbing process. Its key strengths are its import/export tools and the ease of use of its querying tools. Some advanced features were missing from the querying tools such as the ability to specify 'or' operators within the filtering tool (all filter conditions used an implied 'and' operator). In addition, horizontal scroll bars were missing when viewing the tables raw data you were left to clicking a button to scroll horizontally across a table with several columns.

The Google Fusion Tables tools and APIs would have largely been incidental to this project had they not been combined with built-in Google Visualization APIs. This saved a great deal of programming time and as a result several visualizations could be created in the same amount of time it would have taken to program a single visualization from scratch. Besides the built-in geographic visualization tools used to produce the images in this paper, several other interactive visualization tools are available in the Google Fusion Table's visualization tools. These include intensity maps, line graphs, bar graphs, pie graphs, scatter charts, motion clips, timelines, and storylines. Since the data used in this paper was primarily geographic based, the map visualization tools and features were used. However, the other visualization tools made available could prove more useful for other types of data sources. Despite their power and ease of use, the Google Visualizations API features exposed by Google fusion tables did have some shortcomings. In general, most of the features were somewhat in-flexible. For instance, the biggest pain point was the heat map visualization. While this was immensely valuable in showing high density areas of collisions, the colors and sensitivity of the heat map could not be altered. In addition, there was no underlying number that could be attributed to the colors used in the heat map. For instance, when the

heat map displayed read areas there is no indication given as to how dense an area must get before the color red is displayed.

As far as the analysis of IFWIS road kill data itself is concerned, the various images produced for this paper show several ways visualizations can be used to help user comprehension or aid analysis. Visualizations for both interactive and summary uses were produced. Data was organized to show not only the bare location markers on a map, but also show the types of collision by certain species, as well as photos and other information in the pop-up displayed when a particular incident was clicked. Collision occurrence density was visualized according to county and heat map. While the county visualization could prove useful for government, insurance, and other category based analysis, the heat maps proved most useful in showing geographic areas of the State where collisions have occurred most frequently.

The approaches taken in this paper could be considered a starting point for future analysis, various other visualizations could be created from the IFWIS data such as: making the colors on the county visualization relative to county size in square miles, merging traffic accident data per species into visualization (i.e. hitting a Moose is much more dangerous to motorists than hitting a deer), and showing density of collisions per section of highway. Yet, despite the various other visualizations that could be created from this data, the visualizations in this paper revealed several possible trends and patterns that in the least warrant attention or further investigation:

- In general, summer months are the safest driving months as far as wildlife collisions are concerned.

- The Yellowstone area of Idaho is most dangerous in the Summer and Fall.

- Collisions are most likely to happen around border areas between mountains and valleys.

- Wolves appear to have been killed at a rate much higher than would be expected given expert opinions of population and/or behavior.

- Several rare species have been killed on highways including River Otter, Golden Eagle, American Marten, Wolverine, Grizzly Bear, and Mink.

- The majority of wildlife collisions occur in just 7 of Idaho's Counties.

REFERENCES

[1] *Nonfatal Motor-Vehicle Animal Crash--Related Injuries --- United States, 2001--2002.* (2004, August 5). Retrieved March 2, 2011, from CDC MMWR Weekly:
http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5330a1.htm

[2]*Google Fusion Tables API.* (2011). Retrieved 3 2, 2011, from Google Code Labs: http://code.google.com/apis/fusiontables/

[3]*Google Chart Tools / Interactive Charts (aka Visualization API).* (2011). Retrieved 3 3, 2011, from Google Code:
http://code.google.com/apis/visualization/documentation/

[4]*CALPhotos: Animals.* (2009, 8 29). Retrieved 3 2, 2011, from CALPhotos: http://calphotos.berkeley.edu/fauna/

[5] *Visualization: Geomap.* (2011). Retrieved 2 25, 2011, from Google Code: http://code.google.com/apis/visualization/documentation/gallery/geomap.html#Example