Abstracts from the High School Big Data Challenge 2017-2018: Think Global, Act Local with Big Data

Under the Patronage of the Canadian Commission of UNESCO Conference Proceedings
Sponsored by the Royal Bank of Canada

10.17975/sfj-2018-006

Dominik Bednarczyk, Joshua Rosenberg, Tasneem Badshah, Peter Lai, Sean Tao, Kaveeshan Thuraiarajah, Ryan Yu, Ayeze Hassan, Nameera Azim, Andre Cao, Pearl Clam, Parsa Moghaddam, Farbod Naji, Arash Motazedian, Milad Saadati, Nathaniel Chan, Jacky Lee, Tony Liu, Jason Yuen, Seyed Sepehr Seyed Ghasemipour, Shayan Ghaffari, Rishabh Jain, James Kotic, Tony Xu, Shayan Khalili, Jonathan Chiang, Isaac Fung, Ritvik Singh, Gabrielle Terekh, Ashlee Jiang, Daniel Pechersky, Kayley Ting, Serena Perera, Lily Azzopardi, Adi Fishkin, Andrew Schmittat, Jack Zhang, Evan Chandran, Yingyi Liang, Zhami-lya Bilyalova, Haohao Fu, Cheng Guo, Rohit Menon, Jason Arbour, Jordan Juravsky, Shahar Lazarev, Josh Zwiebel, Jonah Garmaise, Ethan Ohayon, Mason Silver, Jonah Belman, Joseph Train, David Roizenman, Seth Damiani, Ronny Rochweg, Katherine Gotovsky, Alain Lou, Arielle Shannon, Jing Yi Wang, Yixuan Chen, Mo Chen, Loredana Cirillo, Haoru Meng, Jamie Birker, Valerie Hermanns, Akera Otto, Olivia Wignall, Aaron Abraham, Kevin Lin

Disclaimer

These abstracts are provided for all student teams that have submitted project reports by January 15, 2017. The STEM Fellowship Journal editorial board has made every effort to ensure proof and English editing of these abstracts in a limited amount of time, and neither organization as a whole or any of its volunteer members can be held accountable for inaccuracies that may have occurred in the abstract publication. Abstracts are published in alphabetical order per school names of the participating teams.

STEM Fellowship's Big Data Challenge is a unique pedagogical experiment, providing an inquiry and learning experience for high school students that, upon equipping them with top-notch analytical tools, tasks them to find hidden patterns and trends in complex scientific data. This year's challenge provided a multidisciplinary competitive opportunity; over a period of three months, teams analysed sustainability data through the prism of computational methods in order to promote and further the UN's 2030 Agenda for Sustainable Development.

Teams applied a variety of perspectives in their projects via the UN's 17 Sustainable Development Goals. The tools used by teams were similarly diverse, ranging from R and Python to SAS Studio and Tableau.

For all the variation between project themes, it remains that all submissions are of incredibly high quality. Every paper is demonstrative of immense creativity and high potential on the respective team's part.

On behalf of STEM Fellowship, I would like to extend my heartfelt congratulations to all students who participated in the challenge, and I wish them all the best for their future endeavours in research and data science. It has been a privilege for us to witness the analytical capabilities of the next generation of students firsthand, and I am certain all entrants will continue to demonstrate excellence in their respective research careers.

Jaweria Qaiser

Director for Data Science Education, STEM Fellowship
Big Data Challenge 2017-2018

A World Split into Red and Blue: A Report on What Influences Belief in Climate Change

Dominik Bednarczyk, Joshua Rosenberg, Tasneem Badshah
The Abelard School

The discussion of climate change has now become a controversial topic. Whether global warming is human-caused is considered a matter of opinion, though there is ample scientific evidence to say the contrary. The research contained within this paper focuses on the correlation between education and opinion on climate change in the United States. Further focus was directed into the issue by contrasting different levels of education, the impact of political affiliation, government financing of initiatives related to climate change, media reportage of climate change, and the influence of location on attitude towards climate change. The goal of this research was to present what impacts belief in climate change, and to what extent. The primary methods of research were the pandas, matplotlib, and seaborn libraries of Python 3 to create and merge data frames from the relevant data we found, and represent correlations visually in plots. It was determined that

high school graduation, overall, has little to no correlation with belief in climate change (varying slightly based on partisan status). However, in 2015, the percent of those with higher-level degrees correlated strongly with belief that climate change is happening. It was hypothesized that this was due to an increase in research on and media coverage of climate change, as well as current events (such as the beginning of the presidential campaign, and the adoption of the Paris climate agreement). In 2016, a sharp divide in opinion was seen based on political party, and much less correlation with education, possibly because of the 2016 presidential election, the limiting of access to government data, and the spread of the idea of “fake news.” This reveals that partisan divide plays a huge role in the conversation around climate change, and is a step that needs to be overcome as individuals move to solve the environmental crisis.

An Investigation into the Relationship between the Extent of Climate Change Research and Climate Change Action in Universities

Peter Lai, Sean Tao, Kaveeshan Thuraijah, Ryan Yu
Bayview Secondary School

Universities and affiliated research institutions produce a significant portion of climate change data. Researching at the bleeding edge of human understanding, they not only provide the data on climate change, but the means to face it. However, although some universities prioritize funding for climate change and preach urgent action and education, do they take measures themselves to reduce their own negative impact?

Various Ontario Universities were analyzed based on publicly accessible data on public grants, total funding, student population, and greenhouse gas (GHG) emissions. The data was plotted using five-year moving averages to reduce local discrepancies. K-means analysis divided the data into four clusters of GHG per capita emitters. It was found that institutions that allocated relatively little funding varied in the per capita GHG emissions.

However, it was discovered that universities who had more research funding allocated to climate change research had consistently lower emissions; all such universities fell into the two lowest emission clusters, and those with the highest funding into the lowest. This seems to suggest that some though not all universities are reducing their footprint regardless of how much they invest into climate change research, yet those who do put an

emphasis on climate change research consistently have lower per capita GHG emissions. This finding adds credibility to the data coming from institutions that invest significantly into climate change research, and is a victory for climate change education.

Urban Planning : Correlation Between the Business Development Areas and Traffic Emissions

Ayeze Hassan, Nameera Azim, Andre Cao, and Pearl Clam
Earl Haig Secondary School

The chief objective of this study is to investigate a major source of air pollution, traffic emissions, in order to deduce a reason for the large amount of vehicle volume in certain areas and to find solutions to aid in reducing it. Data obtained from the Toronto Open Data supported our hypothesis that a staggering number of vehicles make their way towards the business development zones in the Greater Toronto Area every day. This inevitably leads to escalated traffic congestion on the highways and roads heading towards the financial districts, causing those areas, as well as neighbouring residential areas, to have a significantly high amount of air pollutants. For our research, we have decided to focus on nitrogen dioxide (NO₂) as it is one of the dominating air pollutants and has significant health and environmental concerns. Moreover, in order to support this theory, we have generated graphs comparing the volume of vehicles at various intersections spread across the GTA over a period of 8 hours on a specific date with the estimated NO₂ production in surrounding areas, the volume of vehicles at various intersections compared to the business development areas, and the concentration of NO₂ in various areas. A fourth graph demonstrates the positive correlation between NO₂ production between 2011 and 2016, and the number of asthma patients, which supports the fact that NO₂ leads to health hazards and environmental issues, since other factors affecting asthma are fairly constant since the time period is relatively short. Finally, we have proposed solutions which may reduce the need for people to commute such long distances, in turn reducing traffic congestion.

Study of the effect that green house gas emissions have on the rate of death per cancer

Parsa Moghaddam, Farbod Naji, Arash Motazedian, and Milad Saadati

Earl Haig Secondary School

Recently, a number of studies have been done to find the negative impacts of the greenhouse gas emission on different concerning issues like health, economy, and climate change. However, there are an insignificant number of studies that focus on the mutual relation between the rate of deaths per cancer and greenhouse gas emission. Initially, the amount of people affected by cancer in North America and Europe were compared to the amount of greenhouse gases in the atmosphere. Then, new conditions were added to the data to analyze the new outcomes ex. gender, geographical area and the time in which there were infected.

A linear relationship was discovered between these two elements. This relation is best described as the direct negative effect of change in climate to the increase rate of deaths per cancer.

This research can be used to inform people of the danger and the impact of climate change on their health.

Detriments of Climate Change to Ice Wine Vineyards in the Niagara Peninsula Region

Nathaniel Chan, Jacky Lee, Tony Liu, Jason Yuen

Earl Haig Secondary School

This study focuses on the detrimental effects of climate change on the production of ice wine grapes in the Niagara Peninsula. Climate change and rising atmospheric temperatures are negatively affecting crop yields. This publication focuses on the white grape varieties Chardonnay, Vidal, Riesling, and red Cabmet Franc grapes, which are the types of grapes used to make ice wine in the Niagara Peninsula. The process involves harvesting grapes at an optimal temperature between -10°C and -12°C. Farmers will inevitably be forced to relocate if the trends of climate change continue.

Weather station databases were used to determine the changes in temperature and precipitation in the Niagara Peninsula over the last 80 years. Data from vineyards was compiled to reveal the climate conditions suitable for harvesting the grapes used for icewine.

Through this process, optimal temperature and precipitation ranges for icewine were identified. The scatter plots were analyzed and a trend was seen using cross analysis. The time when the region would be unsuitable for grape yields was observed.

An analysis of the results revealed that as the atmospheric temperature rose, the region's effectiveness for ice wine grape growth was hindered. The investigation revealed that by 2030, the Niagara peninsula will not be the optimal location to grow hybrids of white grapes. The region will be unsuitable for the germination of red and white grapes following the year 2097.

The observations have applications for vineyards and winery businesses. The dynamic process of relocating wineries is a logistical setback. As temperatures rise, vineyards will be forced to change their locations, destroying the legacies of vineyards which gives wine their high price. With this concern, many wine companies need to forecast the climate conditions in the near future so that they can plan ahead and transition to more suitable climates, as current vineyards become unusable.

A Study On The Correlation Of Carbon Dioxide Emissions And Poverty

Seyed Sepehr Seyed Ghasemipour, Shayan Ghaffari, Rishabh Jain, James Kosic

Earl Haig Secondary School

In order to explore the correlation between global poverty rates and carbon dioxide emissions, we analyzed datasets from the World Bank Data documenting many important factors relating to energy production. Through the investigation of the amount of CO₂ emissions, percent of the population in extreme poverty, and the amount of renewable energy consumption, we discovered a number of interesting correlations: Countries with higher poverty rates tend to also have higher renewable energy consumptions, while countries with low poverty rates are more reliant on fossil fuels. However, in both cases, large quantities of carbon dioxide are released. Further observations revealed that in poorer countries, in the majority of cases, energy was obtained through biomass due to its cost efficiency and accessibility. Investigations into this matter revealed that preparation and maintenance of biomass leads to the release of large quantities of carbon dioxide, presenting a strong hypothesis for why our initial observations suggested no direct correlation between carbon dioxide emissions and poverty.

A study into the demographics having the greatest carbon footprint

Tony Xu and Shayan Khalili

Earl Haig Secondary School

Greenhouse gas (GHG) emissions have been one of the biggest contributors to climate change and global warming, especially since the start of the industrial era. Studies show that atmospheric carbon dioxide concentrations have risen from 280 ppm to 400 ppm in the past 150 years, and combined with the rises in concentrations of other greenhouse gases, has caused a 0.85°C increase from 1880 to 2012. Although many studies have been done on greenhouse gas emissions and its relationship with climate change, socioeconomic trends in relation with greenhouse gas emissions are not well studied. The aim of this research paper is to determine which demographics produce the most greenhouse gases, and therefore propose policies to minimize the average greenhouse gas emissions per capita. Unlike past studies looking at a single particular factor, this study analyzes 14 factors to compare and contrast their respective relationships with GHG per capita. The 14 factors analyzed were: literacy Rate, GDP per capita, rural population percentage, urban population percentage, adolescent fertility rate (births per 1000 women ages 15 - 19), unemployment percentage, percent of agricultural land, research and development expenditure, renewable energy consumption (% of final energy consumption), food production index, population growth (annual %), mobile cellular subscriptions, air transport freight and forest area (% of total area).

The Effects of Renewable Energy on Canada's Future

Jonathan Chiang, Isaac Fung, Ritvik Singh, Gabrielle Terekh

P.A.C.E. (The Academy for Gifted Children)

It has become increasingly evident that one of the ways to ensure a sustainable future is to invest in sustainable energy resources while phasing out non-renewable sources. Sustainable resources are renewable and therefore release little to no waste into the atmosphere. This paper analyzes the capacity and generation of different sources of power across all provinces and territories in Canada from 2005 to 2015 in order to analyze how Canada can reduce its consumption of non-renewable sources, and how

it can work towards building a sustainable future. The data used was provided by Statistics Canada and it was analyzed using Microsoft Excel. With the use of various graphs, how the energy generation was divided amongst all the provinces and territories was determined, as well as which renewable resources Canada used the most. The data revealed that Canada is slowly increasing its renewable energy consumption while decreasing its non-renewable energy generation. In addition, another major topic that was analyzed were reasons why power plants did not run at full capacity. It was found to be due to two major reasons; maintenance and demand. If long-term energy storing technologies were implemented in the energy grid the need to merely meet demand would be mitigated and thus Canada could generate approximately 97 percent of its energy from renewable sources.

Anthropogenic Climate Change and its Effects on Developing Nations

Ashlee Jiang, Daniel Pechersky, Kayley Ting

P.A.C.E. (The Academy for Gifted Children)

This study investigates natural disaster intensity and its relationship to climate change. Due to increased food and water insecurity, developing countries are more vulnerable to the devastating effects of natural disasters than the countries responsible for the majority of global greenhouse gas emissions. Data was obtained from global natural disaster databases and 6,180 global natural disasters from 2005 to 2012 were analyzed. Python was used to create a dataframe to easily manipulate and investigate relationships between countries' gross domestic products, greenhouse gas emissions, and effects of anthropogenic natural disasters through scatter plots, pie charts and bubble graphs. Analysis of global greenhouse gas emissions data identified several G7 countries as some of the top global emitters. While some developing countries also had high net greenhouse gas emissions, the intensities of their natural disasters were significantly higher. Developing countries had some of the lowest emissions, but alarmingly high fatalities. Recognizing that the adverse effects of climate change are by no means confined to a country's borders, data on disaster intensity sheds light on a pressing issue. The results of this study stress the importance of the global responsibility that all countries have in preventing the devastating effects of climate change on developing nations. Helping nations develop without a reliance on fossil fuels for fuel production, investing in climate

resilience, and lowering our global emissions will be crucial for guiding our steps towards sustainable development.

The Effects of Natural Disaster on First Nations

Serena Perera, Lily Azzopardi, Adi Fishkin, Andrew Schmittat
P.A.C.E. (The Academy for Gifted Children)

The goal of our project is to examine the consequences of climate change on First Nation communities. As aboriginal groups have often been marginalized throughout history, our purpose is to shed light on the specific issues facing natives as climate change is manifesting itself by way of forest fires, floods and harsh winters. This information will allow us to see the relationship between issues that have plagued First Nation communities, such as the need for relocation due to natural disasters and the diminishment of their monetary funds in dealing with these issues. We are using technology to relate the locations between many First Nation communities and the locations recent natural disasters. We can then examine the effects these disasters are having on their respective communities in terms of measures like relocation and fatalities. Ultimately we are looking to show the consequences of climate change on this specific demographic of the Canadian population to better assess the struggles they face and determine how best to serve them.

Efficiency of Clean Energy Investments By Country: a Data Analysis Using Linear Regression and Principal Component Analysis

Jack Zhang, Evan Chandran
Phillips Exeter Academy

Today, improving energy sustainability and investment in renewable energies are the primary methods of combating anthropogenic climate change. However, it is often unclear where the best places to invest in clean energy are. Usually, it requires months of intensive research into the political and economic situations of countries to determine the effect of energy-sector investment in generating electricity and reducing greenhouse gas emissions. However, through an analysis of climate and economic data from the World Bank and Enerdata, which included variables such as greenhouse gas emissions by country and per capita energy production, a general guide to future investment can be developed.

Principal component analysis was applied to explore high-dimensional economic and climate-related data in R, and linear regressions were applied to factors that were found to be related to investment efficiency. It was found that already-sustainable countries benefited less per dollar per capita invested in them $p=2.11e-7$, and that a larger investment per capita reduced the effectiveness of each dollar per capita spent $p=0.0418$. A proposed optimal per-capita investment by country was determined by a mathematically derived equation.

This analysis suggests that to generate more electricity while minimizing greenhouse gas emissions, investing in currently carbon-inefficient countries is preferable over developed countries. However, real-world considerations such as corruption and instability were not included within this quantitative analysis. This means that the analysis does not provide a definitive guide to international energy investment, but instead a general guideline to focus researchers' attentions. These results lay a framework for in-depth analysis on methods to combat the greatest current existential threat to humanity.

A Model of Renewable Energy Generation in the World with Focus on Canada

Yingyi Liang, Zhamilya Bilyalova, Haohao Fu, Cheng Guo
Princeton International School of Mathematics and Science

Renewable energy is becoming a more and more important topic as our natural resources drain. Many social factors are influencing and being influenced by the use of renewable energy, such as economy, education and environment. To explore this, a model was built with selected social factors - education expenditure, GDP, population density, surface area, electricity access and research expenditure - that might influence the use of renewables for each country. Renewable energy was represented by the ratio of electricity generated by renewables to total electricity generation for the countries. Data from World Bank, Government of Canada and Wikipedia were analyzed with R in RStudio, and a statistically significant model was built. The model showed that GDP and education level have a positive relationship with the electricity generation percentage by renewable energy, while population density and electricity access have a negative relationship with the electricity generation percentage by renewable energy. The change of this percentage over time was then looked at more closely for provinces in Canada, and a gap between Canadian provinces with

high generation percentage by renewable energy and provinces with low percentage was noted. Future research could be done on adding more factors to the model, then building and comparing Canada's model with models for the world to see if the same factors are influential.

Lets Dig In: A Study on Crop and Livestock Production and Climate Change in California

Rohit Menon

St. Francis High School

California is the largest agricultural producer of fruits, nuts, vegetables and milk in the United States. The purpose of this study was to examine the effects of lettuce and milk production in the top seven agricultural counties on climate change - average daily maximum temperature (Tmax) and greenhouse gas (GHG) emissions - in California based on data available from several open sources. Python and R were used for data extraction from these sources, and datasets were formulated. Using R and SPSS, regression analyses reveal that in lettuce producing counties, the best combination of indicators to predict GHG emissions are total acreage for leaf lettuce, freshwater withdrawals, and sprinkler irrigation. There were no significant predictors for Tmax. In milk producing counties, the best combination of indicators to predict Tmax is GHG emissions from farm equipment and operations, and the best combination to predict total GHG emissions is the number of cattle head and total freshwater withdrawals for irrigation. There was no significant difference on Tmax or total GHG between lettuce or milk counties during the years 1990-2015. Results suggest that it is important to examine various irrigation practices in these counties to mitigate the rise in daily maximum temperature and total greenhouse gas emissions.

Going Green: Cost and GHG Emissions in the Canadian Electricity Sector

Jason Arbour, Jordan Juravsky, Shahar Lazarev, Josh Zwiebel
TanenbaumCHAT

Electricity generation in Canada contributes significantly to national carbon emissions. In order to meet global greenhouse gas (GHG) emissions targets, Canada must pivot away from generating electricity from fossil fuels such as coal and natural gas and move

towards electricity sources with smaller carbon footprints, such as wind, solar, and hydroelectricity. However, because renewable energy technology grows progressively more cost-effective with time, there is an economic incentive to withhold investment in the short term.

This study determines the optimal time and type of energy to invest in to keep costs at a minimum while still meeting worldwide emissions targets. Our group established an emissions ceiling (the maximum amount of carbon dioxide one can release into the atmosphere while adhering to emissions targets) for the Canadian electricity sector by proportionally scaling down the worldwide GHG cap. Using data from OpenEI, our group used feedforward neural networks to predict the levelized cost of energy (LCOE) for Canada's three largest renewable sources of electricity (wind, solar, and hydroelectricity) and its two biggest fossil fuel sources (coal and natural gas) over time. Canada's future demand for electricity was predicted using the same technique, and with data provided by StatsCan. Our group developed an ideal investment plan by meeting each year's increasing electricity demand with the least expensive generation method of that year. When the emissions ceiling was reached, electricity from fossil fuels would be completely phased out, making the Canadian electricity sector completely carbon-free.

Results demonstrated that wind energy is currently the most cost-effective method of generating electricity (figure 2), and should be invested in immediately to meet growing demand. We predicted that the electricity sector will reach its emissions ceiling by 2033 if one followed the ideal investment plan outlined above. When the ceiling is reached, solar, hydroelectric and wind-based electricity generation will all be more cost-efficient than either coal or natural gas, providing multiple viable replacements for fossil fuels.

The Impact of Climate Change On the Agriculture Sector

Jonah Garmaise, Ethan Ohayon, Mason Silver, Jonah Belman
TanenbaumCHAT

This study deals with the analysis of the future impact of climate change on the agriculture sector. In order to analyze data for this paper, many diverse methods were used. Using tools such as Numbers, complex and carefully created statistical models and data sets were created for the express purpose of analyzing

climate change's effects on the agriculture sector. Agriculture is one of the most important economic sectors in the world and is constantly growing. However there is a new threat on the horizon for agriculture: climate change. Climate change has the potential to wreak havoc on the agriculture sector if not stopped soon. Climate change seems like a relatively new phenomenon. However, climate change has actually been gradually taking place since the industrial revolution, becoming more powerful everyday, and now threatens humanity. Climate change is no longer a thing of the future. If action is not taken against climate change, it is clear that the Earth, and everything living on the Earth, will suffer greatly. Three plans on how to deal with climate change are examined by this study. The first is one where humanity makes no changes to its current ways, the second is the Paris Climate Accords and the third is a plan where temperature increase is limited to 1.5 degrees Celsius above pre-industrial levels. The effects that each plan will have on climate change and the agriculture sector is examined. The conclusion of this study is that the world needs to change its ways in order to avoid climate change and save the agricultural sector.

An Analysis of Historical Global Warming and Social Engagement

Joseph Train, David Roizenman, Seth Damiani, Ronny Rochwerg
TanenbaumCHAT

The goal of this paper is to determine whether there is a correlation between awareness of global warming, and where global warming occurs. This theory is carried out by analyzing maps containing various forms of data that have to do with global warming, such as precipitation and surface temperature, and comparing it with a map of engagement from tweets which mention global warming. This paper found that there is no solid correlation between mentioning global warming in tweets and global warming's effect, although there was a steady increase in both. This is most likely due to Twitter's user base increasing over the years. Therefore it appears that although the effects of global warming have increased, the percentage of people aware of it on major social media sites has not. This then concludes that before trying to find a solution or preventative measure to global warming, an approach must first be made to create awareness for it on social media platforms.

Identifying optimal locations for photovoltaic solar farms in Ontario, Canada

Katherine Gotovsky, Alain Lou, Arielle Shannon, Jing Yi Wang
University of Toronto Schools

This study aimed to identify optimal locations for photovoltaic solar farms in Ontario, Canada. One-hundred and sixty-two Ontarian sites were ranked according to the suitability of their climate conditions, using data on insolation, daylight hours, cloud coverage, humidity, and temperature. The purpose of this study was twofold: one, to help policymakers identify locations for future solar farms, and two, to highlight the strengths and weaknesses of current solar farm locations. The data used in this study was retrieved from the Ontario Climate Normals Database, the Ontario Climate Change Portal, and the NASA Prediction of WorldWide Energy Resource (POWER) database. Using this data, an algorithm was designed to determine the suitability of each municipality's climate conditions for solar farm adoption. It was concluded that the optimal locations for solar farm construction in Ontario were in the regions surrounding the weather stations of Culloden Easey, Delhi, and Hagersville.

Global Warming and Life under the Water

Yixuan Chen, Mo Chen, Loredana Cirillo, Haoru Meng
Villanova College

With the development of society, people have already stepped into a fast-developing industrial age. The use of fossil fuels, coal and petroleum has become a main source of carbon dioxide emission, which is the key of global warming. Carbon dioxide, in recent years, is known as a greenhouse gas which has caused a rise in global temperature. There are many harmful effects global warming would bring to humans. Due to dramatic climate change, it is time to raise people's awareness towards this ecological and social issue. The purpose of this research was to find whether global warming would have any impact on life under the water and how that can be used in future predictions. This research report looked at the water data for Ontario Lake, one of the freshwater lakes among the Great Lakes. The databases of daily water surface temperature, ice concentration, and concentration of chlorophyll absorbed in water have all proven that global warming did have

great influence on lacustrine organisms. Also, according to the trend that was produced from the databases, a rough future trend was able to be produced. At the end, there are also some suggestions on what people can do to reduce environmental impact.

Minimizing the Negative Impacts of Crop Growth on the Environment, and Human Population

Jamie Birker, Valerie Hermanns, Akera Otto, and Olivia Wignall
Villanova College

This research paper was produced for a secondary school level data analysis competition where the theme presented was “think global, act local.” Based upon this objective, focus was directed towards linking the many factors that influence both the environment and human population. The goal was then narrowed to determining ways in which the positive impact of crops can be improved through resource allocation and on the population by way of customized nutritional analysis and optimizing labour efforts. Figures comparing various crop factors, including water usage, yield, harvested area, production, and nutritional value were generated. These figures were subsequently analyzed to produce a set of proposed actions with the potential to maximize the positive influence of the growing agricultural industry on the environment and human population.

Communities and Crime: A Study on Factors Related to Violent Crime Rates in the United States

Aaron Abraham, Kevin Lin
Webber Academy

This study aimed to find the major factors that are closely correlated with violent crime in communities scattered across the United States. A dataset from the UCI Machine Learning Repository was used to gather statistics on potential demographic and socioeconomic factors that could affect violent crime rate. Five models were chosen (linear regression, stepwise regression, relaxed LASSO, random forest, and extreme gradient boosting trees) to predict which variables are the most correlated with violent crime rates. Models achieved RMSE scores of approximately 0.13 and r^2 values of 0.65, indicating reasonable fitness. All models indicated a mix of demographic factors, socioeconomic factors, familial factors and individual factors that

are closely correlated with violent crime rates, including, but not limited to: percentage of kids per two parent households, number of children born to unmarried couples, and percentage of divorced males in communities. Visualizations on important variables indicate that there is correlation to crime rates, which could be traced to competition over resources, psychological harm, and difficult living conditions. Hopefully, these results can be used by law enforcement agencies to accurately tackle crime and reduce violent crime rates.