2021 Highschool Big Data Challenge: Paving the path to true equality and equal access in education

Under the patronage of Canadian Commission for UNESCO

10.17975/sfj-2020-012

STEM Fellowship's High School Big Data Challenge is an inquiry-driven experiential learning program that provides students an opportunity to learn and apply the fundamentals of data science – a crucial skill set for a young researcher in the digital age – through independent research projects. The COVID-19 pandemic disrupted high school education, at the same time creating a “fertile ground” for interdisciplinary, student-driven STEM education.

This year, we invited students to explore issues of *Equality and Equal Access in Education* and to suggest their own evidence-based solutions, using Open Data and the principles of Open Science. Students explored many topics, ranging from using machine learning to find hidden socioeconomic factors in access to education, to the efficacy of various modes of instruction.

We developed in-depth learning modules designed to lead the student from zero-knowledge to an elementary working proficiency in data science. The students learn a broad range of data analytics tools and programming languages which are useful for uncovering hidden patterns, trends in structured and unstructured data. Some of the tools the students learnt and used includes Python, R, LaTeX, and machine learning.

On behalf of the STEM Fellowship, we extend our sincere congratulations to all students who participated in the challenge, and wish them the best for their future endeavours. We want to express our appreciation to all the mentors and volunteers. This program would not be possible without patronage of CC UNESCO and generous support of our sponsors: RBC Future Launch, Let's Talk Science, Digital Science, Kimberly Foundation, SCWST, CISCO Academy, Canadian Science Publishing, Faculty of Science UofC.

It has been a privilege for us to witness the analytical capabilities of the next generation of students firsthand, and we are certain all entrants will continue to demonstrate excellence in their respective careers.

Anish Verma

Chief Data Officer

Dr. Sacha Noukhovitch

Founder, President and Editor-in-Chief

Disclaimer

These abstracts are provided for all student teams that have submitted project reports by January 23, 2021. 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 by title.

A Case for Technology - Spending and Equity

Andrew Butler

Ascension Collegiate, Bay Roberts, Newfoundland

The presence of digital learning technology is often seen as the hallmark of the modern classroom. An increasing number of classrooms worldwide are integrating these technologies into their lessons. A notable indicator of the presence of digital learning technology is educational spending. This paper involves quantitative analyses of potential issues of equity regarding educational spending, particularly on classroom technology, and students' academic benefits from it. Notable disparities were found in standardized testing between students from countries with high per capita educational expenditure and those who did not. Even in those countries that possess large educational spending, differences in performance still existed between sexes despite improvements in the mitigation of those differences. Analysis and visualization of diverse data frames from intergovernmental organizations through the use of the Python programming language that confirms the existence of disparities in access to digital learning technology and resulting academic performances.

A Correlative Analysis of Educator Identified Factors and Global Student Assessment Scores

Natalie Jarycki

Oakridge Secondary School, London, Ontario

Effective system structure and fiscal spending in education are ongoing goals of educators and governments across the globe. This study analyzed whether there is a correlation between specific variables identified on the Teaching and Learning International Survey (TALIS) and student average scores on the Programme for International Student Assessment (PISA). Eleven variables felt to capture teacher ability, morale, autonomy, and performance review tools and implications were analyzed. The goal was to determine predictive power of whether improvements in certain educator identified factors could improve student education outcomes. The study reflected 600,000 students

and 275,000 teachers and school leaders in 44 countries that participated in both the 2018 TALIS and 2018 PISA. It was hypothesized that nations with low TALIS scores would also demonstrate low PISA scores reflecting either direct teacher or indirect education system factors. From an analysis using Microsoft Excel, Python's Pandas library, and Statistical Analysis Software (SAS), only the incorporation of student external results into teacher performance reviews showed a statistically significant correlation, explaining 10% of national PISA score variability. The study results support the ongoing administration of regional standardized testing as a factor in improving education effectiveness. They also suggest that the educational impact of other variables such as social factors would warrant correlative study to help guide impactful fiscal spending.

A Review of the Association of State Testing and District Expenditures in California from 2014 to 2019

Brian Hsiao, Kelly Hu, Olivia Chen

Torrey Pines High School, San Diego, California

As high school students in California who have lived in other states such as New York, it is evident that California under-spends on K-12 education. Statistically, California has large class sizes and less support for students because there are lot of students, but not enough teachers due to the higher-than-average pay for teachers in the high cost state. By analyzing different school districts in the state of California, we were able to find that our prediction was correct. To help increase higher test scores, governments should be paying more expenditures towards schools in poverty areas. For this study, we looked into specific districts throughout California to see if there were any trends between the district's expenditures per student and state test scores. By using data spanning from 2014-2019, we evaluated the data in Jupyter Notebook using python as well as Google Sheets to better visualize our graphs that showed California state testing result trends throughout the years. From the graphs, it clearly showed most districts having an increase in high scores and a decrease in low scores because of the increased expenditures each year. However, some schools had a decrease in the percentage of high scores and an increase in the percentage of lower scores

even though districts increased their expenditures, which prompted us to dive deeper into this finding.

A Study Into How Teaching Practices Affect Learning Outcomes

Stephanie Yi Fei Lu

Maple High School, Ontario

The aim of this study was to determine which teaching-related indicators have the strongest effect on literacy outcomes, and therefore to propose policies to improve the quality of teacher training to ensure equitable education for all. Data from 48 countries and sub-national entities (2018) was analyzed using a multiple linear regression model through the use of a number of Python libraries from Anaconda. It was concluded through the analysis of regression coefficients that the average proportion of class time spent on classroom management, the average proportion of class time spent teaching and learning, the proportion of teachers who feel professional development (PD) in the last 12 months had a positive impact on their teaching practices, teachers who agree that teachers and students usually get on well with each other, teachers for whom student behaviour and classroom management was included in their formal training, and teachers who felt well prepared for teaching in a multicultural or multilingual setting were the indicators that had the most significant impact on predicting literacy outcomes in the model constructed. This suggests that policies focusing on opportunities for PD, training in classroom management, and the accessibility of additional teaching resources would contribute the most effectively to improving equity in education.

A Study of How Money Invested in Secondary Students' Education Impact the Quality of Education in the Math and Science Fields

Raymond Ma, Rachel Dietrich, Ruby Miller, Tenzin Dhonyoe

Runnymede Collegiate Institute, Toronto, Ontario

One of the most important endeavours that every country

seeks to establish is a high quality education system. Success within education is built of many factors, one of the greatest being the funding that goes toward it. The purpose of this study is to analyze how the money invested in secondary schools impacts the students' quality of education in the math and science fields. Inequalities arise in students' quality of education due to the different amount of money countries invest on students' learning. We hypothesized that more money invested in schools would increase the quality of secondary students' education. To prove this we analyzed data from around the world and combined different data sets to create graphs. To represent the quality of education, we measured average test scores in math and science. We mainly used Google Colab, where we programmed in Python to analyze our data and construct our graphs. After the analysis, it was clear that the percent of GDP per capita which goes toward student-to-teacher ratios and teacher salaries has a strong effect on the average test scores, and therefore the quality of education. After our research, we conclude that money is one of the main factors affecting students' quality of education, therefore governments worldwide should invest more in secondary school math and science education programs.

Academic-Related Stress: Student Well-being's Contributions To Optimized Academic Performance

Brooklynn Bennitt, Harry Nguyen, Matthew Braho, Nathan Wong

Abbey Park High School, Ontario

In a world where most of our success derives from our education, we ask, how can stress impact equity among students? How is student well-being affected by academic stress? For adolescents, stress originating from school is quite common, and it can motivate students to succeed. However, there is a point when it is too much to achieve any productivity. Stress can result in physical reactions because of stress hormones such as cortisol. They occur when the body perceives a threat, which in this case is schoolwork. There is bound to be some form of stress for students, but how much is too much? With this, we look to find what the optimal amount of stress is that can maximize academic success,

therefore making the education more equitable from student to student. In the remaining duration of the report, we look to analyze data surrounding student well-being, how each factor is affected by stress, and ultimately how stress relates to academic performance.

An Investigation into the Socioeconomic Factors Correlated to PISA Reading Scores

Orchee Haque, Sophie Hoyer, Stephanie Huynh, Sraddha Uppili

Westmount Charter School, Calgary, Alberta

In this investigation, our aim was to utilize indicators of the Human Development Index (HDI) to determine the socioeconomic factors most correlated to a country's literacy as measured by its score on the Programme for International Assessment (PISA). We examined the relationship between national PISA reading test scores and a country's mean years of schooling, primary school enrollment, secondary school enrollment, GDP, GDP per capita, government spending, life expectancy, infant mortality, and total population. Triannual data spanning from 2000 to 2018 was gathered for over 80 countries and analyzed by year using multiple regression models. Our results were interpreted using R² and Pearson Correlation coefficient values. From these results, we concluded that infant mortality, mean years of schooling, and life expectancy exhibit the strongest correlations with PISA test scores, suggesting that the robustness of a country's healthcare system is a strong indicator of its quality of education.

Analyzing Gender Disparity in Education Through Government Expenditures

Jack Wallace, Ryan Rakusin, Tarik Irshad, Christopher Law

Crescent School, Toronto, Ontario

In order to determine the gender disparity in global education, we compared spending habits between developed and developing countries and analyzed primary, secondary, and tertiary education data from the World Bank. Our findings suggest that there isn't sufficient evidence to

support the claim that government expenditures towards education reduces gender disparity at any of the three main levels of schooling. On a global scale, our primary and secondary graphs demonstrate very slight skewness towards male enrollment while our tertiary graph shows that there is a skew towards female enrollment, at least for countries where data was available. We concluded that women who are given the opportunity to become educated are more likely to continue their education and further themselves than men in similar situations. As for those countries that still have a majority male enrollment, to eliminate gender disparity they must do more than simply invest more funds in their education system.

Application of Multivariate Analysis in Identifying Factors Influencing Education Equity

Andrew Mao, Binder Nie, Nicholas Dong

Earl Haig Secondary School, Toronto, Ontario

Achieving greater equity in education is essential for improving student and teacher well-being and ensuring social fairness. It requires a better understanding of the challenges faced by individual students and teachers and the problems at a societal level. One of the best ways to identify and assess key factors that need to change to achieve education equity is to analyze the importance of hypothesized factors such as percent female teachers, total expenditure per student, and home life. We built extreme gradient boosting (XGBoost) regression models and asymptotic regression models to examine the impact of economic, social, medical, and educational factors on the Programme for International Student Assessment (PISA) scores, an indicator of the quality of an education system, using data from the Organization for Economic Cooperation and Development (OECD) and Our World in Data. The girl child marriage rate, percentage of female teachers, adolescents with computer and internet access at home, adolescents with access to personal space, and total expenditure on educational institutions per full-time student were found to have the most significant impact on education quality. This study suggests that data-driven decision-making can quickly identify pain points in the education systems of different regions. Therefore, we recommend decision-makers use state-of-art techniques to

analyze data before developing policies and plans.

Assessing Factors Influencing Literacy Rate to Isolate Variables Affecting Equity in Education in 2017

Ariana Pritha, Allison Iun, Sreejita Das, Sathya Mathiyalahgan

Colonel By Secondary School, Ottawa, Ontario

Education is hindered by various social factors and the most prominent factors were ones that contributed to the United Nations SDG goals, such as, gender equality, and water rights. As such, there must be a strong determination in the field of scientific research to determine flaws that create a detriment to a nation's advancement. Thus, this report begins its analysis using data science. Using multi-output linear regression models to discover correlations between literacy rate and variables such as health expenditure, seats in parliament held by women, and many more, one can begin to investigate pertinent issues directly correlated to lower literacy rates. Furthermore, this study aims to provide a basis of understanding on how education is connected to a vast variety of variables of both demographic and geographic means to determine issues of greatest concern. Results in certain variables that were deemed to require emphasis due to their correlation weight index resulting in being >0.40 (indicating a moderate strength in correlation) were domestic health expenditure, health expenditure as proportion to GDP, share of female youth not in education, alongside water and sanitation. Research into causations of correlations with these variables and literacy rates will allow public officials to focus on both local and national issues with greatest impact, and consequently will encourage new policies to be implemented to achieve the sustainable development goals.

Ayacucho: Peru's Model Province to Equal Access to Quality Education

Camila Figueroa, Micaela Sousa and Reyna Dai

Canaway Academy, Vaughan, Ontario

Our aim was to determine factors predicting the inequality

to access quality education in Peru and to identify the demographics and geographical areas being affected the most by education inequity. The relationship between accurate government spending and the satisfactory performance of students is key to determine the level of quality education provided to students in each province. We found that the quantity of money allocated to each province is not the most impactful factor, but the way in which each regional government allocates the resources depending on the geographical location and socioeconomic needs. We based our analysis on the relationship between the percentage of successful performance on the ECE (Census Evaluation of Students), with government spending on education per capita, the number of students enrolled in the public schooling system, the average population, geographical area, geopolitical issues, and rate of illiteracy. Our research parameters were the 26 provinces of Peru within the course of 4 years from 2015 to 2018 given that ECE was cancelled. In order to analyze and organize the data, we used multiple linear regression and double entry tables on excel. With this information we came to the conclusion that an accurate government spending related to the population of each province and the efficient allocation of money for education and basic needs, had the most significant impact on the quality of education and student success.

Can Budget Distributions Indicate School District Wealth?

David Zhang

Fremont High School, Sunnyvale, California

As the wealth gap and achievement gap continue to grow ever wider in the United States, more work needs to be done on understanding trends at the intersection of these two disparities. This study examines whether the wealth of a school district can be determined purely through budget distributions; that is, whether the proportion of budget allocated for certain areas and/or the proportion of revenue generated from certain sources can provide insight into how much total revenue is received by the school district. These possibilities are explored in two ways using nationwide financial data provided by the U.S. Department of Education's National Center for Education Statistics. Budget distribution

data is analyzed for trends indicative of such prospects and is also used to develop machine learning models able to classify school districts into wealth categories. Statistical analysis demonstrates that more affluent school districts spend more on programs that contribute to a more supportive education, spend less on executive administration, receive disproportionately high funding from property taxes, and receive disproportionately low funding from other programs such as school nutrition initiatives, among other trends. Furthermore, the machine learning models offer strong performance: training XGBoost on both revenue and expenditure variables results in an accuracy of 99.147% and a log loss of 0.0253; training on only revenue variables results in an accuracy of 95.630% and a log loss of 0.1236; finally, training on only expenditure variables results in an accuracy of 72.003% and a log loss of 0.6862.

Corporal Punishment as a Predictor for Children's Mastery of Foundational Learning Skills

Kuljeet Chohan, Ivan Hu, Alix Dorfsman, Alisha Malik

Westmount Charter School, Calgary, Alberta

The purpose of this study was to examine the relationship between the use of violent discipline by caretakers and a child's mastery of various learning skills. Datasets from UNICEF Multiple Indicator Cluster Survey 6 (MICS6) responses were used to compare the prevalence of violence discipline with various parameters assessing the mastery of foundational learning skills in 21 regions. Based on previous research, it was predicted that an increase in the prevalence in violent discipline would be associated with a decrease in foundational learning scores. This prediction was supported as it was found that the percent prevalence of violent discipline in each region had moderately strong negative correlations with various foundational learning skills. It was found that the level of difficulty of different numeracy outcomes did not have a significant effect on their degree of correlation with the prevalence of violent discipline and violent disciplines' correlations with numeracy skills were found to be slightly stronger than with literacy skills. Through statistical analysis, was also identified that regions with full

prohibitions on corporal punishment had, on average, a lower prevalence of moderate to severe violent discipline than those with partial or no prohibitions. These results lend to the importance of decreasing the prevalence of corporal punishment through legislation that fully prohibits corporal punishment, in order to seek better education for all with respect to the UN Sustainable Development Goal.

Correlation Between Funding and Educational Disparity

Kevin Guo, Andrew Lu, Jim Guo, William Zhang

St. George's School, Vancouver, British Columbia

Social class has always been a strong indicator of a child's educational success. It is one of the many factors which may positively or negatively impact educational outcomes for that child. This is significant because the development of strong cognitive skills are essential, especially later in life. A study by the Economic Policy Institute in 2017 found that while the economic inequalities have grown, the gap in academic performance of sample kindergarten classes has stayed the same. To combat the prominent issues, certain solutions, such as increased aid and comprehensive support, have been implemented, but to limited measurable success to date. The Foundation Skills Assessment (FSA) is the annual British Columbian provincial assessment given to students in grades 4 and 7 to gauge their performance. It is one of the many assessments to collect data and quantify the performance of students across the province. The Fraser Institute, a research body independent of the Ministry of Education, uses this data to rank primary and secondary schools. This data, among others, will be compared to the local average housing cost to determine trends in, or correlations between, performance and housing price. There is a clear correlation between the average home value and FSA scores. Although specific values of scores and real estate prices differ from province to province, the general trend suggests that home prices are a prominent factor correlating with the success of students in a certain region. Other trends in the data sets show that population and housing prices have an inverse correlation, which is similar when comparing population and FSA scores as well. By equally distributing the funding across Canada, the average quality of education for students would improve,

thus reducing inequalities in education.

Economic cost and employment potentials for STEM and Non-STEM undergraduate majors in Canada

Luyuan Ge, Fengyuan Zhang

Bloor Collegiate Institute, Toronto, Ontario

Choosing suitable undergraduate major for each high school students is critical for their future career life. In this study, our aim is to provide the latest information from the view of economic cost and employment potentials for various undergraduate majors. Data were obtained from Canada Government public open database and secondary analysis and description was performed. The results showed that whether Canadian students (CS) or international students (IS), dentistry, medicine and pharmacy were the most expensive majors. The tuition fees of law for CS and veterinary medicine for IS are also expensive. In terms of the fastest growth of tuition fee, nursing, veterinary medicine and education were on the top rank among CS, while the category of life science and technologies, the category of math, computer and information science and social science were on the top of rank among IS. After graduation, employment potentials showed the superior performance in the category of architecture, engineering and related technologies, the category of mathematics, computer and information science and the category of science, technology, engineering and mathematics (STEM) which were reflected by higher proportion of employed full time, permanent opportunity and estimated gross annual earnings. In conclusion, STEM based majors had higher economic cost during program study but superior employment potential after graduation. While non-STEM majors had the opposite performance compared to STEM majors. High school students should weigh the pros and cons of majors and make the optimal judgement in choosing undergraduate majors.

Educational Inequality Between Urban and Rural Areas

Joshua Lakdawala, Yash Jagirdar

Erindale Secondary School, Ontario

Within education, inequalities can significantly impact students and their ability to succeed in the future. A major inequality present globally is the disparity between urban and rural education. This paper analyzes the correlation between education rurality and academic performance within Ontario schools through the lens of UNESCO's 4th Sustainable Development Goal. The purpose of this research paper is to provide sufficient data and solutions to aid government officials to improve schooling and reduce inequalities within the education system. Data utilized in this paper includes 2020 EQAO data from all schools all across Ontario, analyzing location, socioeconomic status, enrolment, and academic achievement. School board budget data was also analyzed in order to determine funding disparities and amount spent per student. Microsoft Excel and Tableau were used to extract and visualize relevant data as scatter plots, quartile distributions, bar charts, and bubble maps. This paper concludes that students within rural areas tend to score lower academically compared to urban students. Factors that have been attributed to this inequality are the distance to an urban area, enrolment in a school, spending per student and socioeconomic status. In order to mitigate this, it is proposed that school funding and transportation accessibility in rural areas is increased. A mobile education resource app has also been developed to help disadvantaged students improve academic performance. Overall, this inequality is clearly a pressing issue, and through the solutions proposed can hopefully be mitigated for future generations.

Effect of Low Income on the Sleep of Students from Low Income Families

Christopher Cao, Mindy Kim, Josh Bowden, Steve Bae

Ernest Manning High School, Calgary, Alberta

The purpose of this study is to investigate whether or not a student's financial status affects their sleep and consequently, their school performance. It is also to investigate the relationship between the effects of inadequate sleep on students and student mental health. It is known that sleep is crucial for the lives of students, but its steady decline has led to severe health problems with dangerous effects on the

future of students. This is especially true for students low income families, as they have a larger workload than their peers. Using data sets from the County Health Rankings and Roadmaps provided by the University of Wisconsin Population Health Institute, it can be compared to the data with the US 2013-14 State and National Estimations from the Civil Rights Data Collection by using SciKit, a machine learning library for the Python programming language. By using linear regression, we found that counties reporting a high percentage of people sleeping less than seven hours on average typically have higher reports of both frequent mental and physical distress along with a higher percentage of student suspension. We also found that median household income played a significant role in sleep, as lower income families typically had a higher rate of insufficient sleep. Our research leads us to believe that although the lives of students are affected by various factors, sleep plays a primary role in the quality of life for all students and its increasing negligence plays a significant role in the growing issue of mental and physical health for students, something especially true for students in low income families, who are forced to surrender sleep to make ends meet.

Effects of pre-primary education on learning outcomes

Eun Gu Kang, Jennifer Zhu, and Maggie Guo

Earl Haig Secondary School, Toronto, Ontario

This paper explores the effects of enrollment and government expenditures in pre-primary education on student performance in secondary school. Using data from Our World in Data and the Organisation for Economic Co-operation and Development (OECD) database, average results from the Programme for International Student Assessment (PISA) in Sweden, Norway, Denmark, France, and Poland were compared to national expenditure and enrollment in pre primary education. Since PISA assesses the academic performance of 15 year olds, past expenditure and enrollment rates were used from the time period that students taking the tests were enrolled in pre-primary school as per the International Standard Classification of Education (ISCED) classification of pre-primary education. The report revealed that neither enrollment nor expenditure had a strong impact

on PISA scores. Instead, it appears that the relationship between enrollment and expenditure had a greater impact than each individually on student performance.

Equity in Education in Ontario

Avril Alphonse, Julianne Hu, Julianne Faith Domingo, Jericho Marc Mendoza

The Early College at Guilford, Greensboro, North Carolina

The objective of the research was to devise a solution for equity in education in Ontario schools during the pandemic. The lives of many were impacted by the major global event, COVID-19, which led to the research of the topic. To carry out this study, information from credible sources such as Statistics Canada was used. The knowledge of computer science was used to code a program that efficiently organizes collected data sets. Key findings were focused on the resources Ontario students had easy access to which included the amount of available devices in a typical household and accessibility to a stable Internet connection. Due to virtual learning, students are in need of a device for themselves solely. Around 58.4% of families who had stable Internet connection had less than one device per household member. Therefore, a possible solution that aids in increasing equity in education is ensuring students have access to novel academic items necessary for better learning.

Equity in Education - Relationship Between Gross Domestic Product and International PISA Test Scores

Amin Fahiminia, Jonathan Manzara, Jason Manzara, Bentley Chai

Westmount Charter School, Calgary, Alberta

The purpose of this study is to determine the relationship between a country's Gross Domestic Product (GDP) and the education of its citizens. GDP is used as a quantifiable measure of market value and may play a crucial role in the funding of education and how it may impact the education of its citizens. This study uses scores from the OECD's Programme for International Student Assessment (PISA). These datasets

were then filtered to only include countries with both listed GDP and PISA exam scores. They were then compared using linear regression, where a list of each country's GDP (logged to the base 10) is directly compared to PISA scores in the subjects of reading, mathematics and science. Based on the data that was received from our code, there is evidence to suggest that GDP may have an impact on PISA test scores. As the GDP of the countries rose in our analysis, the results of PISA scores increased accordingly and flattened off at around 40000 US Dollars Per Capita. Further research and experimentation into this topic is advised to imply causation between the two variables. Factors other than GDP may have played a role in impacting the PISA scores, especially at the aforementioned GDP threshold. The findings in this observational study should be used to better tackle problems pertaining to equity in education in third-world countries across the globe.

Equity in Education: Three-factor and Longitudinal analysis on World Inequality Database on Education (WIDE)

Mandi Jin

Rideau Park School, Calgary, Alberta

This study was conducted to find a more equitable education system for all students no matter their conditions of ethnicity, social status, gender or even troubles arisen as a result of the Covid-19 pandemic. To add to this, the effects of major factors that play an critical role on indicators of equity in education were analyzed. A data set from World Inequality Database on Education (WIDE) was used to gather statistics to further estimate the inequality gap for different indicators and countries. The upper secondary completion rate was chosen to explore inequity in worldwide education and several stacked bar plots were drawn to describe the differences among demographics and incomes of countries. The learning achievement in reading (upper secondary) was chosen to explore inequity in education in Canada. A three-factor analysis model was built to study the effects of gender, location, wealth and their interaction on learning achievement in reading. After model selection and model diagnosis, an additive model was chosen as the final model and it proved that there resided significant main effects. An analysis of variance (ANOVA) table also proved further

confirmation of this conclusion. In addition, A longitudinal model was built to explore whether observations of learning achievement in reading varied among the different years with four different levels as covariate variables. Three models related to repeated measures were built and the final model was chosen based on AIC value. As a result, it provides a clear indication that the government should take the responsibility and well needed actions to eliminate inequality in education for many disparities that exist within comparisons between groups such as gender, rural and urban areas, and social status of families.

Examining the Effects of Government Funding and Staff Compensation on Public School Completion Rates

Allan Cao, Ziyang Li, Sarah Zuo

Westmount Charter School, Calgary, Alberta

Education, an essential part of developing a nation's quality of life, has been targeted as an area of focus by the United Nations Sustainable Development Goals. Specifically, Goal 4 aims to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all". [6] An essential part of reaching this goal globally will be through the efficient use of funds to create optimal educational systems. This study investigates the link between various facets of a nation's public education expenditure and its school completion rates. Using two variables provided from the World Bank's EduStats dataset, teaching staff compensation as a percentage of total expenditure in public institutions and initial government funding per student as a percentage of GDP per capita, the correlation to education completion rates was analysed utilizing data provided by UNESCO's Global SDG Indicator Database. Scatterplots were plotted for each subdivision of the two variables: primary, lower secondary, and upper secondary schooling. Statistics for all datasets were obtained. The study found no linear correlation between a nation's teaching staff compensation as a percentage of total expenditure in public institutions at any level of schooling in relation to its school completion rate. Weak to moderate positive correlations between initial government funding per student as a percentage of GDP per capita at the primary and lower secondary levels in relation

to school completion rates were found to be statistically significant. At the upper secondary level, the results showed a weak and statistically insignificant positive correlation. This led to the conclusion that an increase in government funding for education, particularly at lower age levels, benefits holistic nationwide educational achievement. Further research into the effects of teacher skill on both teacher salary and educational attainment would be welcome.

Factors Affecting Education Around The Globe

Casey Fong, Bilal Siddiqi, Michael Cassidy

Leaside High School, Toronto, Ontario

The objective of this study stems from the 4th UNESCO sustainable development goal "Equity In Education." This goal entails using data science and data science methodology in order to help achieve equal education for people across the world. This study in particular focuses on young children in first world countries because these are the minds that will control the entire world in the coming decades. By improving education with young children, not only are we improving the quality of education, we are also helping develop any field that these millions of children decide to pursue in their professional career. The data sets in this study have two main purposes; determining the countries with the best educational systems, and determining which factors those countries have in common. The "best" countries were determined by which percentile of a given data set the country lies in, and assigning a point value. Then, the point value of each country was averaged from all the data sets they had values for, and any countries with insufficient data were omitted from the running. After the top countries were found, they were compared to determine which data sets had the strongest impact on creating a strong educational system.

How Geographical Factors Influence Student Test Scores

Aneya Kazmi, Wendy Chen, Alice Du, Kelly Zhao

Earl Haig Secondary School, Toronto, Ontario

This paper examines the correlation between Quality of

Life indicators in regards to the quality of education, and its impact on standardized testing results. Utilizing data from past PISA test scores, average marks and outcomes to compare the correlation of the average Gross Domestic Product per Capita, Government Expenditure, the HDI and the Life expectancy, from 5 different Continents was investigated. We used Pandas, Matplotlib, Seaborn and Scipy in Python to present our analysis. As a result of investigating the causes and effects of education inconsistency, there was a significant amount of data that inclined towards a conclusion that the Quality of Life indicator had a relationship with Education, and that the problem of education disparity was a direct result of increasing poverty within certain nations.

How Natural Disasters Affect Funding for Education and Job Opportunities

Cindy Tsou, Serena Li, Veronique Wong, Valarie Wong

Earl Haig Secondary School, Toronto, Ontario

This report aims at discussing the relationship between frequency of natural disasters and distribution of funding. We will dive into the issue of whether the money spent on education is able to maintain a balance. Utilizing data from government Canada and reliable source such as Kaggle, we examine the correlation between natural disasters and educational funding. Our analysis are performed using Excel and Google spreadsheet. It is worth to observe that global warming increase the numbers of natural disasters and indirectly affect the economy in each provinces.

How Public School Funding Will Decide the Success of Our Future Leaders and the Bright Minds of Tomorrow

Deniz Jasarbasic, Eric Karpovits, Ana Balteanu, Aidan Bergin

Lisgar Collegiate Institute, Ottawa, Ontario

Education is the foundation of an equal-opportunity society and is essential to the development and success of young people. Across the world, public secondary education

is funded using various methods of distribution; however, some funding systems may be more effective than others. This study will analyze different education funding systems and their effectiveness in delivering students the best experience and overall success. Our objective is to perform a comparative analysis between the equal distribution method of public school funding in Canada and the varying distribution method of public school funding in the United States of America (US) using open-source education data gathered from both countries between 2007 and 2017. We analyzed the data using multiple regression models with the goal of identifying advantages and disadvantages to both systems. If this impartiality gave every student a fair chance to succeed or given the data, it led to a worse overall quality in education, experience, and success during and after their secondary education. From there, using an integrative thinking method, we will determine some new ideas that will improve the funding system from a student-success and impartiality point of view, providing novel perspectives about public school funding that can be scaled at various levels of education.

How School Survey Data Tracks Academic Performance

Allen Pei, George Wang, Edward Xiao, Daniel Yang

University of Toronto Schools, Toronto, Ontario

Equity in education continues to be a pressing issue for administrators and the students they serve. Infact, governments strive to find the best factors and data which separate "turnaround schools" from their mundane cousins. In this study, we investigate the hypothesis that school surveys are an underutilized tool which provides key insights about school conditions and learning environments. We examine student achievement and school-wide survey data gathered by the New York City Department of Education for the 2018-19 school year. To be specific, we compare a carefully calibrated academic achievement score to the percentage of positive or negative responses to survey questions. This is a significant departure from prior research, which has focused on covering each category of education indicators independently. We used a variety of techniques including linear regression and an XGBoost model to measure and

display correlation. They demonstrated that bullying and harassment were the most significant indicators of academic success, while the best performing questions addressed to students had higher correlations than those addressed to teachers, and especially parents.

Inequity in Education: the Correlation Between Neighbourhood and School Performance in Ontario

Shuyu van Kerkwijk, Florence Ho, Hibah Hussain, Grace Ko

University of Toronto Schools, Toronto, Ontario

An effective education system should strive to provide an equitable education for all students, regardless of background. To assess the success of the Ontario education system in carrying out this function, we examined the correlation between a school's performance on standardized tests and its neighbourhood factors (e.g. economics, demographics, geographies) for Ontario elementary and secondary public schools. The EQAO is a standardized test in Ontario that tests students on multiple subjects. We gathered data from diverse sources, including school EQAO results over multiple grades and years, and census data from Statistics Canada. We analyzed the data using multiple regression models, spatial autocorrelation tests, and hierarchical regression. We found that neighbourhood factors can explain around 40% of the variance in school performance on the EQAO, with the percentage of students whose parents have some university education being the foremost factor. Moreover, the degree of this correlation remains relatively flat across different grades, suggesting that schools are failing to reduce the education disparity between students in socioeconomically disadvantaged and advantaged neighbourhoods. We conclude, tentatively, that the Ontario education system has achieved equality in education, but not yet equity.

Investigating the Correlation Between Wealth, Academic Enjoyment, and Happiness

Alicia Bremer, Angelina Chen, Isabella Laing, Matt Mrzljak

St. Thomas More Collegiate, Burnaby, British Columbia

Socio-economic status can influence many aspects of one's life. For example, school performance is impacted by the amount of resources available in one's home, school and country. Differences in wealth affect students' quality of schooling, creating inequity in education. This study investigates whether students interest in and enjoyment of certain subjects is related to one's financial status or access to funding. Our paper investigates the effect of money on student interest and happiness in academia, furthering understanding of how to work towards equity in education in relation to student enjoyment of learning and well-being. We defined a null hypothesis, that increased interest or enjoyment in school does not correlate with wealth. We performed linear regression on data from PISA tests conducted on 15-year-old students in reading, math and science from 2000 to 2018. The resulted in negative R2 values, which suggests that there is no correlation between socio-economic status and increased interest or enjoyment in school. We also displayed this in several scatter plots. In this report, we reveal the impacts of our findings and suggest further research that could be done into improving student enjoyment and using government funding more efficiently.

Investigating the Relationship between Poverty and Educational Success

Albert Choi, Jazli Muhammad, Bobby Tang

Westmount Charter School, Calgary, Alberta

Knowledge is one of the most important resources an individual and society can possess. In parts of the world, knowledge is obtained through a formal education via an educational system divided into several categories (pre-primary, primary, secondary, and tertiary). The cost of formal public education in the majority of countries contributes to a portion of countries national wealth, often measured by its gross domestic product per capita. The costs of education at renowned universities in countries with GDP per capita such as Canada or the United States are known to be considerably high. The cost of education gives a good indication of the quality of education worldwide, where higher levels of education may lead to higher national prosperity. This study explores the OECD PISA education score of 62 countries,

particularly its relationship with GDP per capita, primary school enrollment, and class teacher to student ratios. PISA score data was gathered using data collected by OECD statistics. GDP per capita, primary school enrollment rates, and teacher to student ratio data were collected from the World Bank Group. Each factor was separately compared to the OECD PISA education scores of countries, and a linear regression was performed on each of them. The results of the linear regressions indicated that countries with higher PISA education scores had higher GDP per capita, indicating that countries with higher primary school enrollment rates had a higher quality of education. The results led to the conclusion that countries of higher quality education lead to a higher GDP per capita, of which the significance of teacher to student ratios is valuable.

Investigation on the Relationship Between Oil and Gas and Influenza in Alberta

Erin Chen, Chloe Liang, Manal Kamran, Jakub Gierus

Weber Academy, Calgary, Alberta

A correlation often exists between the economy of a region and the health of its citizens. Therefore, this study aims to investigate the relationship between the oil and gas prices and production in Alberta (a prominent part of the province's economy), with influenza cases in the region from 2014 to 2019. An extensive weekly report of influenza cases was extracted from the FluWatch, and multiple data sets detailing oil and natural gas prices and production was collected from the Alberta Government Economic Dashboard. Using Microsoft Excel, correlation values were calculated and ranked. Visual representations of the data sets were also created in the form of scatter plots and bar line graphs. The results indicated that there is no significant long term relationship between influenza positivity rates and oil and gas factors. Within certain years, particularly 2015 to 2016, and 2018 to 2019, stronger correlations were found between Flu A rates and oil and gas prices. However, due to the limited data set and time frame investigated, further research would be needed to reach a definitive conclusion. The results of this investigation could highlight the possibility of underlying economic factors contributing to the rise of influenza, which

can also lead to further research in examining the relationship between the health sector and oil and gas industry across the globe.

Is Free Education Worth it? Modelling the Effects of Legally Guaranteeing Free Education on Student Outcomes

Ahmed Al-Hoda, Alka Ukrani, Jonathan Barazzutti, Paul Milan-Diaz

Westmount Charter School, Calgary, Alberta and Sir John A. Macdonald Secondary School, Waterloo, Ontario

Ensuring inclusive and equitable education is Sustainable Development Goal 4 in the United Nations' (UN) Sustainable Development Goals (SDGs) which are to be achieved by the year 2030. This paper aims to examine guaranteeing free primary and secondary education as a solution to SDG 4, by looking at how free education affects the quality and inequality of an educational system measured using a wide range of educational indicators. We calculated a variety of relative index scores that include those indicators to measure the quality of educational systems around the world. We then used these index scores within 16 different multiple regression models to quantify the effects that free education has on student outcomes. Our results suggest that overall, free education has positive effects on student outcomes in terms of reducing educational inequality in addition to improving the average student outcomes, which suggests that free education is a viable solution to SDG 4.

Limitations To Indigenous Youth Education

Rebecca DiFilippo, Alejandra Hernandez, Aiden Robinson

St Ignatius of Loyola Catholic Secondary School, Oakville, Ontario

Our project aims to determine factors that influence post-secondary opportunities for indigenous youth in Canada. There have been previous studies examining how indigenous youth are often underprivileged, but our focus is on the factors hindering their ability to attend university/college. We

analyzed the percentage of indigenous youth in university, percentage of indigenous youth in high school, number of local universities in dense indigenous areas, and costs for students who travel out of their province for university. We gathered population and educational statistics in Canada spanning from 2015-2019 and found two prominent factors inhibiting their success.

Multivariate Analysis of United States Public Education Data to Optimize the Classroom Environment

Ammar Vora, Nishan Soni, Vishwanath Wimalasena, Faiz Hanafi

Sir Winston Churchill High School, Calgary Alberta and Robert Thirsk High School, Calgary Alberta

Optimizing our classrooms to be as successful as possible is key to creating more equity in education. Therefore, the purpose of this investigation is to analyze which factors influence student performance, and by how much. This will give our education systems a list of variables they can change to improve student performance, thereby producing a more equitable educational environment. To accomplish this task, economic, demographic, and educational data from the US National Centre of Educational Statistics (NCES) was used to analyze factors that may affect student performance. Within the analysis, a plethora of materials and methods were used. These included the Python programming language, Jupyter Notebook and multiple statistical tests used to interpret the results. These results ranged from finding strong relationships, to weak correlations, to no appositions at all between the factors influencing education. For example, increased class sizes, educational appropriations, teacher salaries, and the percentage of students with internet access were all found to be influential factors in not only the academic success of students, but also their educational well-being. Each one of these findings can be interpreted in different ways, and therefore must be applied in ways unique to themselves. Innovative solutions are required for this type of research, and future studies can build upon the work laid out in this paper by applying the same methodologies and using the same formula to discover how else we may change the societal modus vivendi to build a better world for all.

Predictive Modeling and Visualization of Equity in K-12 Education using Machine Learning

Roma Desai, Kushi Jakati, Keval Shah, Alexander Creech

The Early College at Guilford, Greensboro, North Carolina

Equity in education is a critical issue facing our society as outlined by the United Nations Goal #4. Access to education can be impeded upon by various factors including: gender, race, and socioeconomic background. This study evaluates the impact of socioeconomic background, minority status, and behavioral aspects on the standardized testing scores in K-12 schools in North Carolina, United States (USA). Data sets were collected and processed from National and State databases including the North Carolina Department of Public Instruction and the National Center for Education Statistics. Data visualization was performed using dashboards for 270 schools throughout 100 counties in North Carolina across elementary, middle and high school levels. Statistical analysis including the Tukey's test was implemented to evaluate differences among different school types (Title I, Title III and Regular). Factor-screening analysis was conducted using ANOVA to evaluate the significant factors affecting student scores. Machine learning was utilized to develop predictive models using six different supervised learning algorithms. The Support Vector Machine algorithm had the highest accuracy (around 90%) to predict SAT scores based on socioeconomic, minority and behavioral input factors. Intervention strategies were recommended based on findings from the predictive models. This research lays the foundation for developing longitudinal programming that can positively impact educational outcomes (scores) of students across elementary, middle and high school levels. In addition, this study focuses on expanding access to education and improving test scores in North Carolina and other developing countries throughout the world.

Teacher Participation and Student Equity: Should teachers have greater agency in large-scale decisions?

Maitri Shah, Richa Srivastava, William Zhou

Westmount Charter School, Calgary, Alberta

In the past decades, technological and living standards have risen across the world, buoyed by increasingly educated workforces. As education improves, it becomes important to understand how different educational factors impact equity. Today, international equity metrics still show serious gaps between learners belonging to different ethnicities, genders, and wealth brackets. Among developing nations, these disparities are especially severe. This paper provides a methodology that objectively examines the impacts of teacher involvement on student equity. This study examines the relations between inequality and in-school factors. The methodology analyzes two sets of linkages. The first is the one between wealth-based inequality and various factors from PISA; the second is the connection between teacher participation and agency — as measured by the Teacher Participation in Decision-Making Index (TCHPARTI) — and inequality. The reason teacher participation was chosen in particular and analyzed to greater depth is because it is a factor that impacts all schools around the world. Data from OECD's PISA surveys and the World Inequality Database on Education and analyzed. Overall, the results of the study show a correlation between the agency given to individual teachers and a reduction in mean inequality. Based on the results, it is recommended that teachers be given an opportunity to increase participation in large-scale decision-making processes as this would help reduce educational inequality due to wealth. Through these developments, teachers would also be able to promote smaller class sizes and the implementation of special needs programs in schools, both of which are factors that have been shown to decrease inequality in education.

The Correlation Between Economic Status and Quality of Education

Isaac Bahler, Samantha Chong, Nanya Anyaduba, Nier Kamel

Lindsay Thurber Composite High School, Red Deer, Alberta

The goal of this report is to determine the link between economic status and level of education obtained. Two different data sets were used to examine this relationship. A combination of worldwide and national data was considered

in order to obtain a global perspective. Due to lack of access to Canadian data, supplemental analysis using national data from the United States was conducted. Data was collected from multiple sources, with a portion being collected using a python program that automatically collected the data from a government website and the rest being manually collected from provided data sets. The data was cleaned using google sheets to determine trends, create graphs, and make other visuals. The data demonstrated a correlation between economic status and educational quality. Based on the data the more financially secure your family is the better your education will be. This demonstrates a fundamental problem with equity within education. Even in affluent zones such as the United States, this inequity persists leading to unequal academic outcomes. It represents one of the underlying problems that remain within educational systems throughout the world and efforts must be made to correct the situation..

The Impact of the Student Nutrition Program on Academic Achievement in Ontario

Cristian Bogdan, Harshith Duba, Nathan Pfeffer, Nada Salem

Earl Haig Secondary School, Toronto, Ontario

This paper investigates the impact of the Student Nutrition Program (SNP) on academic achievement in Ontario. Using the 2015-2016 Canadian Community Health Survey (CCHS) and Ontario student demographic data sets, we compared nutritional indicators and academic performance in Ontario's 36 Public Health Units (PHU). Our data analysis was performed using the NumPy, Pandas, pyKML, Shapely, and Matplotlib Python libraries. According to our analysis, there is no significant relationship between the student nutrition program and academic achievement in Ontario schools. Further research found that educational attainment of parents and household income had a more significant impact in other countries around the world. However, some research has found a positive relationship between their own SNP and student achievement, suggesting Ontario's SNP may be ineffective. To better understand the shortcomings of the current SNP, there must be a greater emphasis on collecting qualitative, rather than quantitative, data on the SNP and conducting further research into household food insecurity.

The SNP must re-evaluate its eligibility requirements to accommodate for schools with lower-income students who need it most.

The Impact of Teacher's Background on Student Performance

Adam Omarali, Mars Xiang, Matton Xia, Timothy Wang

Crescent School, Toronto, Ontario

In her 1954 essay *The Crisis in Education*, Hannah Arendt says, "Education is the point at which we decide whether we love the world enough to assume responsibility for it." While education has been much more widespread and available in the 20th Century, it is important to focus on all aspects of learning. The push for Global Education goes beyond just being able to access education. Instead, the United Nations and much of the world must focus on the equity and quality of every student's future. Improving the quality of education around the world is not a simple task. Thus, one of the most significant areas that Team TAMM decided to focus on is Student Performance impacted by the number of years teaching, experience, and salary of a teacher; as well as the student's socioeconomic status impact on their performance. We hypothesize that, with a closer lens, being put on the quality of education around the world will make a profound impact on students. We obtained the statistics on education using available OECD, World Bank, and government data of 52 states and 51 socioeconomic groups. We organized the data into 4 significant representations which included looking at SES vs student math and reading scores, a correlation between 10 indicators, and year over year percent change. The results indicate the constantly progressing decline in standards throughout each country's education system and highlight the areas that need to be fixed. The implication is that although teacher salaries are increasing, the quality at which students are receiving their education is not.

The Role of Government Expenditure in Equitable Education

Jatin Mehta, Philip D'Costa, Keval Patel

Erindale Secondary School, Mississauga, Ontario

In today's world, formal education is considered a prerequisite for tapping into one's potential and developing young people into financially-stable and successful individuals. However, many student around the world are unable to partake and excel in education as there exists a disparity in education between males and females. Given that there is such a large inequity in education, our team decided to examine the role of government expenditure in equitable education. The aim of this research study is to analyse the impact of government expenditure in education on various aspects of education, namely student participation and student outcomes. To do so, a variety of data from the OECD and UNICEF were used. Using linear regression models, correlation analysis, and statistical distributions, we concluded that increasing government expenditure correlates to an increase in equitable education as it resulted in better harmonized tests scores, and out-of-school rates for secondary school students for both genders. However, there is also a correlation between government expenditure and gross domestic product (GDP) per capita, suggesting that countries that earn more have a better opportunity at increasing equity in education.

The Trend of Quantitative and Qualitative Investment in Education for OECD Countries

Tianyi Xu

Hillfield Strathallan College, Hamilton, Ontario

Quantity and quality of education are often seen as factors that lead to educational development in countries. However, our knowledge of the relationship between quality and quantity of education is limited. Our aim is to examine the trend of quantity and quality of education throughout the years. We analyzed statistical data organized by the Organization for Economic Co-operation and Development. Through our analysis of the trend in both the quantity and quality of education, we believe that investment for qualitative

education plays a more important role in developed countries while investment for quantitative education plays a more important role in developing countries.

Using Feature Importance Methods to Identify the Primary Factors Contributing to Education Index

Samuel Lu, Evan Skrukwa, and Shahryar Wasif

Westmount Charter School, Calgary, Alberta

This study aims to use data to work towards achieving Sustainable Development Goal 4 (SDG 4): ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. To do this, we must first consider the factors contributing to education quality and inequity and focus our attention and resources accordingly. In working towards this goal, we determined if there is a significant relationship between certain indicators (out of all the indicators for SDG 1 & 4 from the UNESCO Institute for Statistics) and the education indexes of countries and ranked indicators by importance. We used 6 methods for evaluating feature importance and selected 20 indicators that have the greatest relationship with education index for each method. The top indicators selected from these methods had 5/6 indicators being selected as capturing linear and non-linear relationships between the education index and the predictors. These results are significant as we now have a better idea as to how attention and resources should be best used to provide positive results in order to achieve SDG 4 most efficiently.

Using Machine Learning to Investigate the Impact of Income, Inequality, and Infrastructure on Education Around the World

Vikram Arora, Darian Ellis, Andrew Kang, Andy Tang

Hillfield Strathallan College, Hamilton, Ontario

Education is the key that will allow many other United Nations Sustainable Development Goals (SDGs) to be achieved. When children are able to access quality education,

they can break from the cycle of poverty, reducing different inequalities around the world. This paper investigates different income, inequality, and infrastructure indices and their effects on the United Nations education index (EI). Using human development data obtained from the United Nations Development Programme, machine learning models were trained to analyze the relationships between education index, mean years of schooling, and a variety of different indicators. Future indicators were forecasted using linear regression and double exponential smoothing. Initial analysis indicates that the income index has a strong positive correlation to EI, while the coefficient of human inequality (CHI) and gender inequality index (GII) have strong negative correlations to EI. Upon examining developmental trends of EI between 2010 and 2018 and retro-extending different indicators 20 years further back to 1990, extensive gaps were observed between countries in different human development index (HDI) tiers. The most concerning findings are the significantly slower annual growths in EI and other factors strongly correlated to EI in countries within the low HDI tier, especially in comparison to countries within the medium HDI tier. This results in the continued widening of already substantial education and wealth gaps. Countries with a higher HDI typically have access to more resources, allowing them to put more resources into education. These countries are also able to invest more in education-related infrastructures, such as electricity in schools and access to clean drinking water for students. However, countries with lower HDIs and higher inequality indices tend to have relatively fewer resources, as well as a much smaller percentage of youth attending school. Without foreign aid, these countries will likely continue to lag behind the rest of the countries in education moving into the future.

do so, gross agricultural sales were compared with county graduation rates in over 700 counties across the United States. Python was used to streamline the data collection process and analyze the data due to the large volume of CSV files. Some states were left out of this study due to unavailable, missing, or inconsistent data, but states from all parts of the country were used to create a representative sample. We hypothesized that higher agricultural prevalence may correlate with a decrease in high school graduation. Our study shows no significant correlation between graduation rates and the agricultural yield of the counties. Considering more indicators and using other countries may be beneficial in order to display the relationship between the degree of focus placed on the agricultural sector of the economy and the level of educational attainment in countries with less developed agricultural industries.

Using U.S. agricultural data and graduation rates to identify factors in accessibility of education

Tim Gubski, Parthiv Maneesh, Martin Bardhi

Westmount Charter School, Calgary, Alberta

This study explores the relationship between the degree of agricultural focus in a community and the level of educational attainment for the individuals involved. To