**Predicting Global Video Game Sales: Analyzing Genre, Publisher, Platform, Region Influences**

Project Ethics, Impact, and Sharing Report

Tanner Patrom

**Project Ethics**

An ethics study is merited to understand the broader implications that a research project can have on society. Ethical conversations range from the effect that the project can have on society to the concerns raised from the collection of the data.

**Data collection:**

Data collection came in an ethically sourced manner. Gathered from vgchartz.com, and published on Kaggle, the data collection methodology of vgchartz states “*on the software front we now only record official shipment/sales data, where such data is made available by developers and publishers*”, confirming that there are no considerations to be made about violating any ethical standards.

**Analysis Techniques:**

The choice of analysis techniques were linear regression and random forest models. This choice was guided by their suitability for the analysis goals and their ability to provide interpretable results for the audience. These models are widely recognized within the data science community for their robustness and have been selected to ensure a thorough approach to understanding the factors influencing video game sales. Care was taken to ensure that the model assumptions and limitations were clearly communicated to avoid misinterpretation of the results. Data preprocessing was also handled with care. Variables with missing values were dropped and platforms were combined to a single creator to normalize the data to enhance the fairness of the model.

**Communication of Results:**

Efforts were made to communicate the results of this study responsibly and with a focus on clarity and accuracy to effectively represent the results to an audience. Figures and graphs have been designed to be clear and informative, providing comparisons and insights into the data, and model results, with little confusion. Conclusions were drawn to reflect the data analysis accurately, without overstepping the bounds of the evidence.

**Impact of the study and results**

**Academic Impact:**

The findings from this project would contribute to the academic field of predictive analytics and economic modeling within the video game industry. Within the digital entertainment sector, a deeper understanding of consumer behavior and market preference could be had by leveraging such statistical models to predict a video games sales based on certain factors. The methodologies applied, particularly in terms of model selection and data preprocessing, can be adopted by other researchers looking to come to similar conclusions in other industries. The approach would bode well in transferring to other consumer behavior observations.

**Local Impact:**

This study could potentially influence video game developers in the regions studied, particularly North America, Japan, and Europe, by providing them with insights into which factors most significantly affect game sales within their markets. For instance, if certain genres or types of gameplay are found to resonate more effectively with local consumers, developers can tailor products to better meet these preferences.

**Broader Impacts**

On a broader scale, the implications of this research extend beyond the video game industry, sectors such as digital marketing and consumer analytics could find influence. Understanding the predictors of video game sales can offer insights into other digital media trends, video games are an enormous industry with many different entertainment domains tying into it. Companies could better strategize to take advantage of these findings.

**Plan for sharing findings**

To effectively target an audience of which the findings in this project would benefit, I found the “Journal of Business Research” to be a good fit. An excerpt from the journals aims and scope says “[JBR] aims to publish research that is rigorous, relevant, and potentially impactful. Recognizing the intricate relationships between the many areas of business activity, JBR examines a wide variety of business decision contexts, processes, and activities, developing insights that are meaningful for theory, practice, and/or society at large.” This project blended data analytics and business research hopefully leading to better decisions made by organizations.

An appropriate conference, where discussing these findings would be of benefit to the audience, would be the Game Developers Conference. Organizing most years since 1988, this conference attracts a wide range of professionals from across the video game industry, including developers, marketers, and researchers. Insights from my study could prove beneficial as individuals attending would be directly involved in that industry.

The source code to this project will be made available on Github under the following repository https://github.com/PhysicalBit7/Intro-to-Data-Science. The repository has multiple directories. The “Data” directory houses the data used in the study and the “src” directory holds all the code used to complete the study.