

AN ANALYSIS ON GALACTIC WELL-BEING

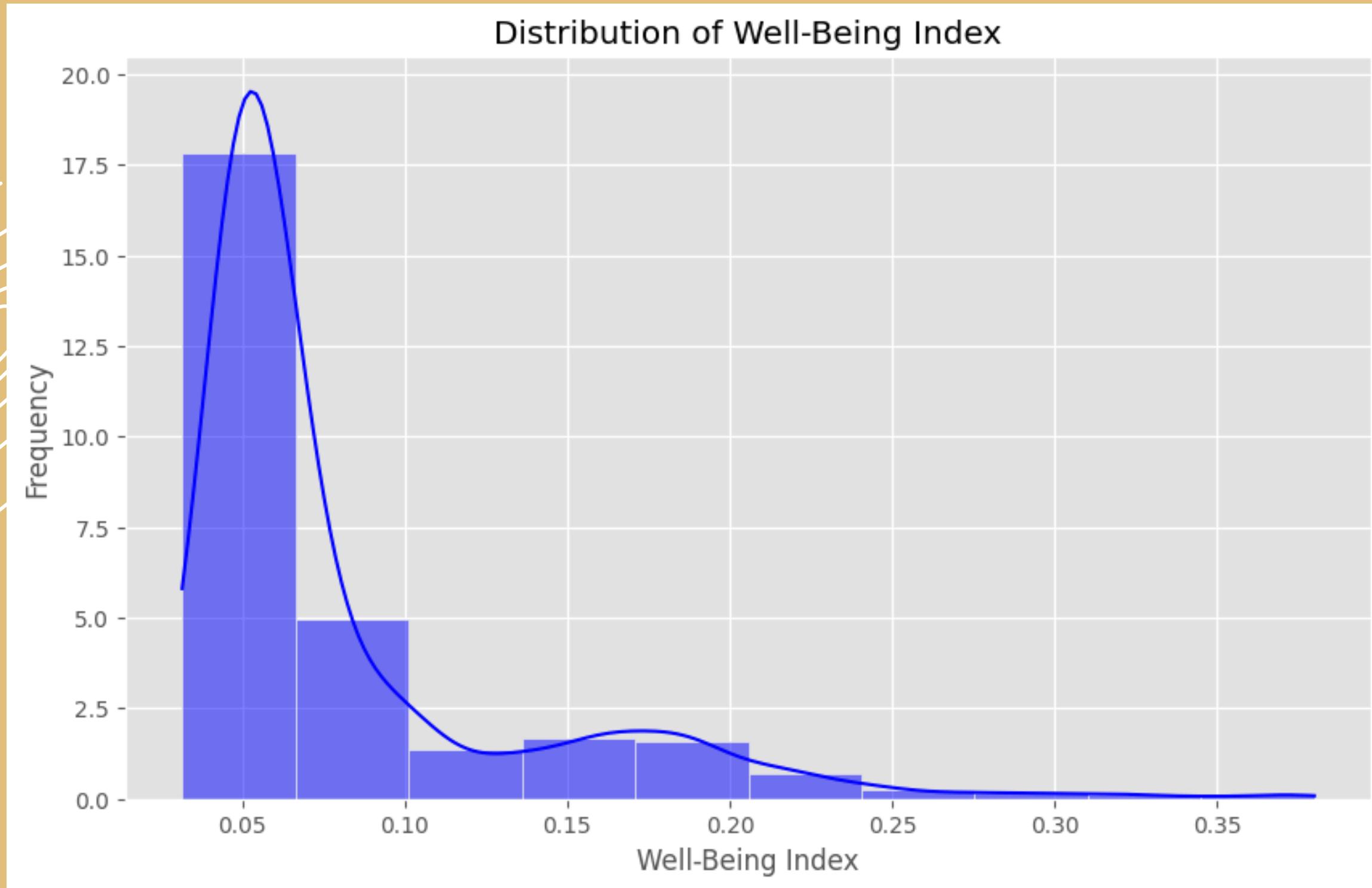
The aim of this analysis was:

- To determine the demographic and socio-economic variables that influence the Well-Being Index of each galaxy. These variables are such as: existence expectancy years, income indices, education years (actual and expected), population, mortality, unemployment rate and gender inequality indices among others to form 80 possible factors.
- To make predictions on the well-being of the new galaxy members provided given the same demographic and socio-economic factors provided.

Initial data handling:

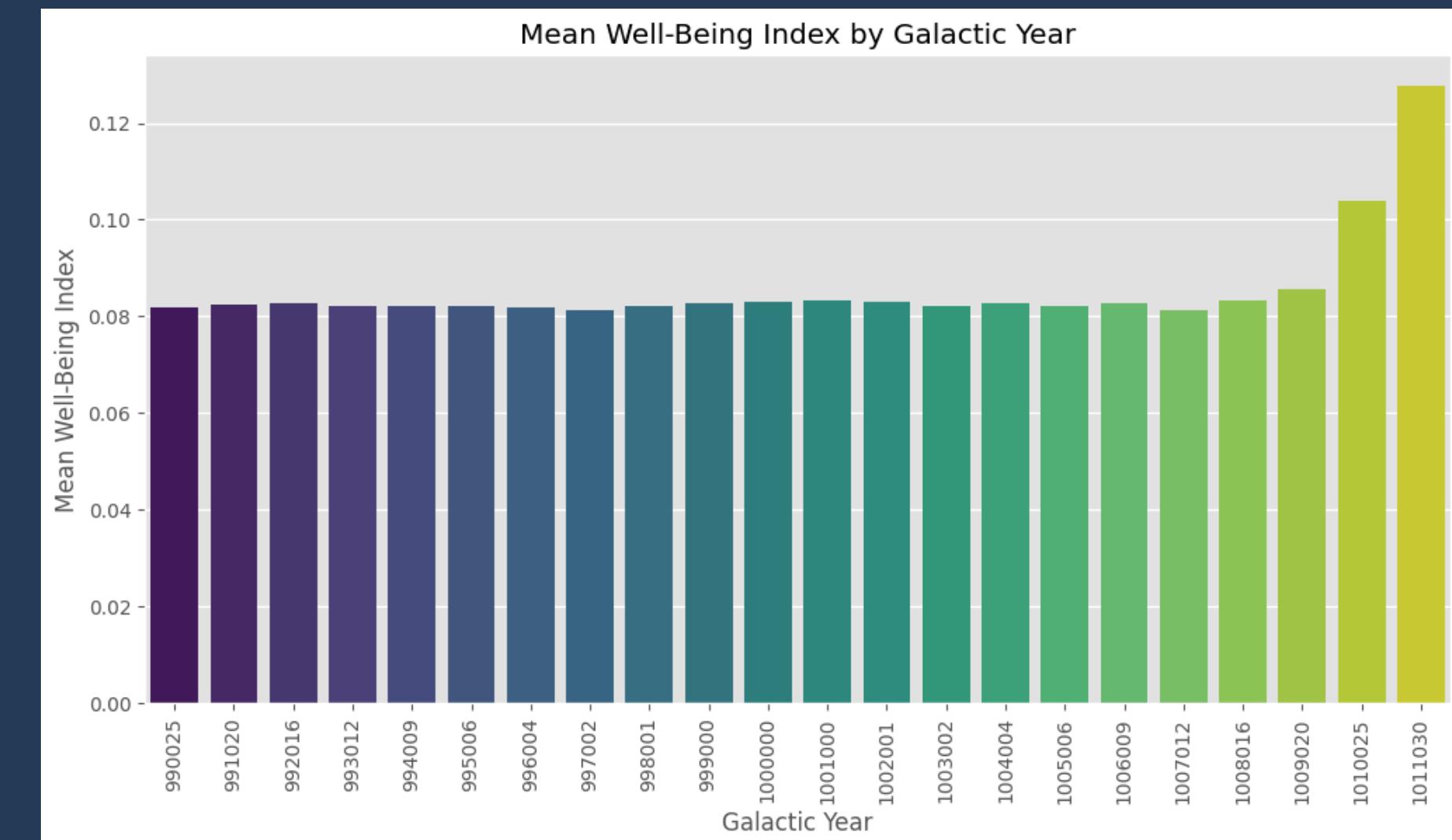
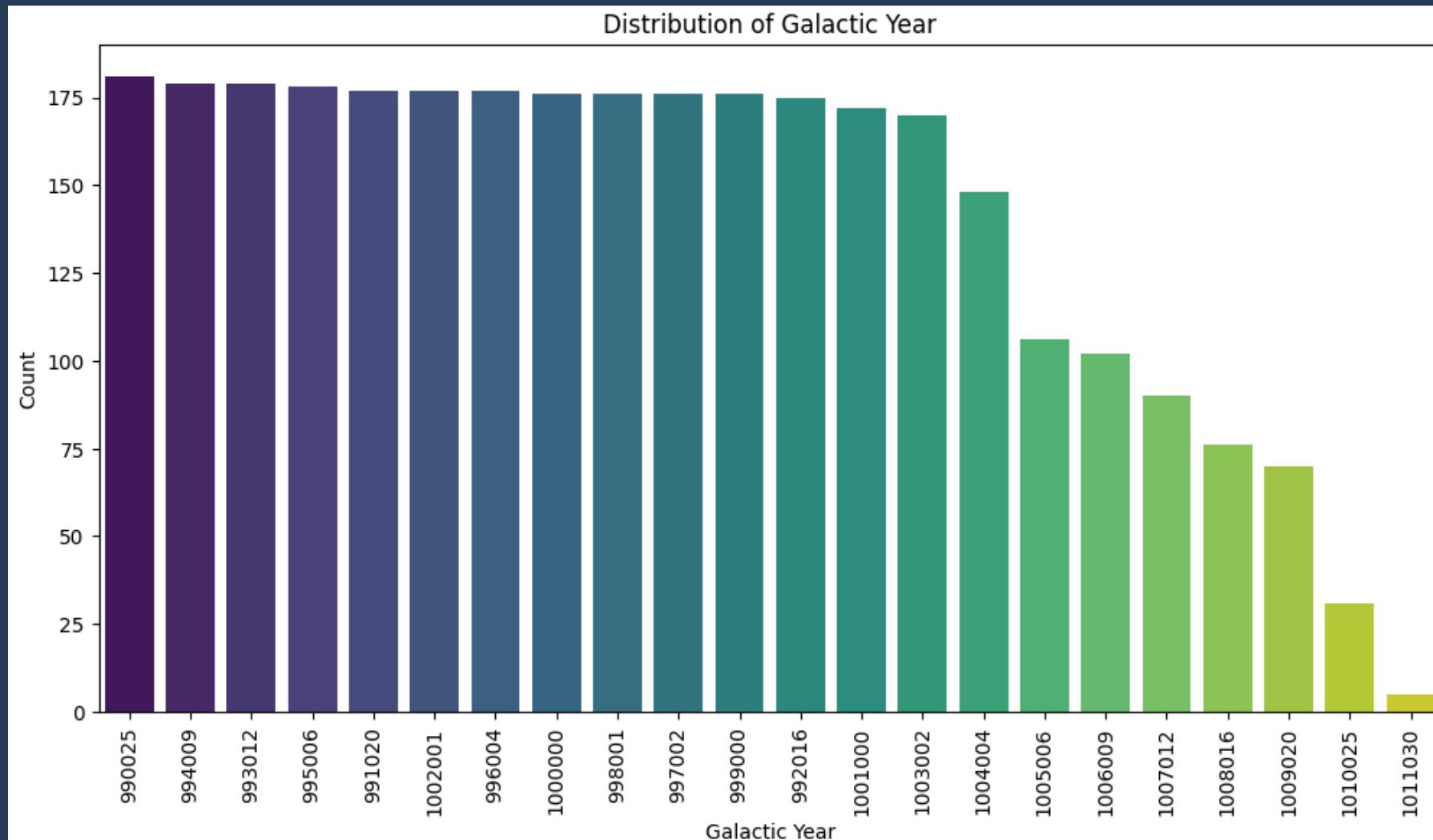
- From the 80 features, we narrowed down to 10 features that had sufficient data (at least 70% filled) to make meaningful analysis. This excludes the ID and galaxy features that were descriptive.
- We established that each ID was unique and that there were no duplicates.
- Filled in the missing values of the remaining columns, if any, with an appropriate measure to ensure completeness.

THE WELL-BEING INDEX



- The Well-Being Index was most frequently recorded as being in the 0.05 to 0.10 range. There were very few observations with an index above 0.25.
- The Well Being Index had a moderately positive correlation to all other features apart from itself, ID, galactic year and intergalactic development index rank
- The Intergalactic Development Index Rank was the highest correlated to the well-being index though the relationship is negative. The lower the rank (numerical not literal), the higher the Well-Being Index.

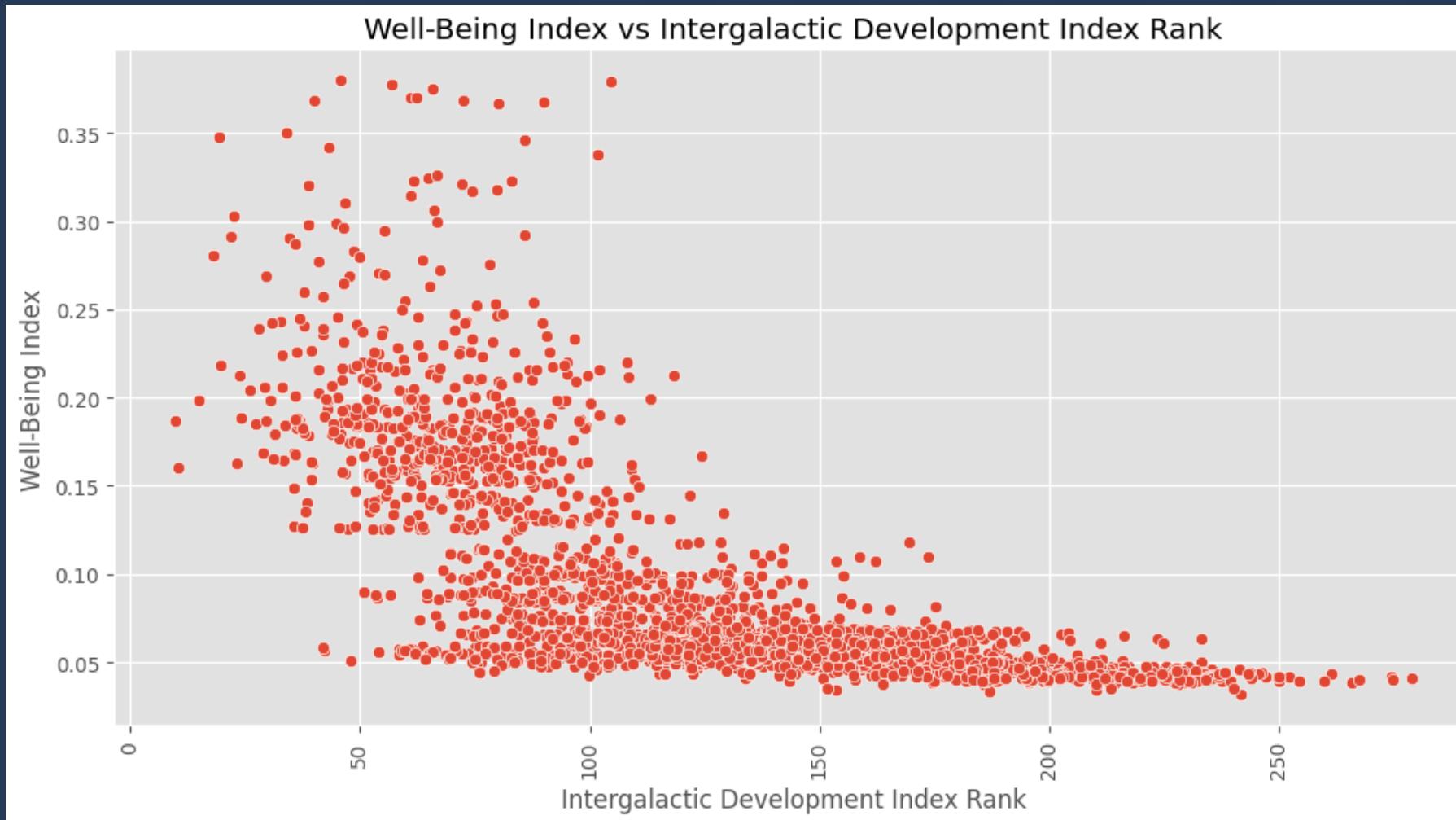
GALACTIC YEARS



- The data had 3097 individual galaxy members over formed over 22 years.
- The galactic year with the highest formations was 181 observations in 990025 and the least formations being 5 in 1011030.

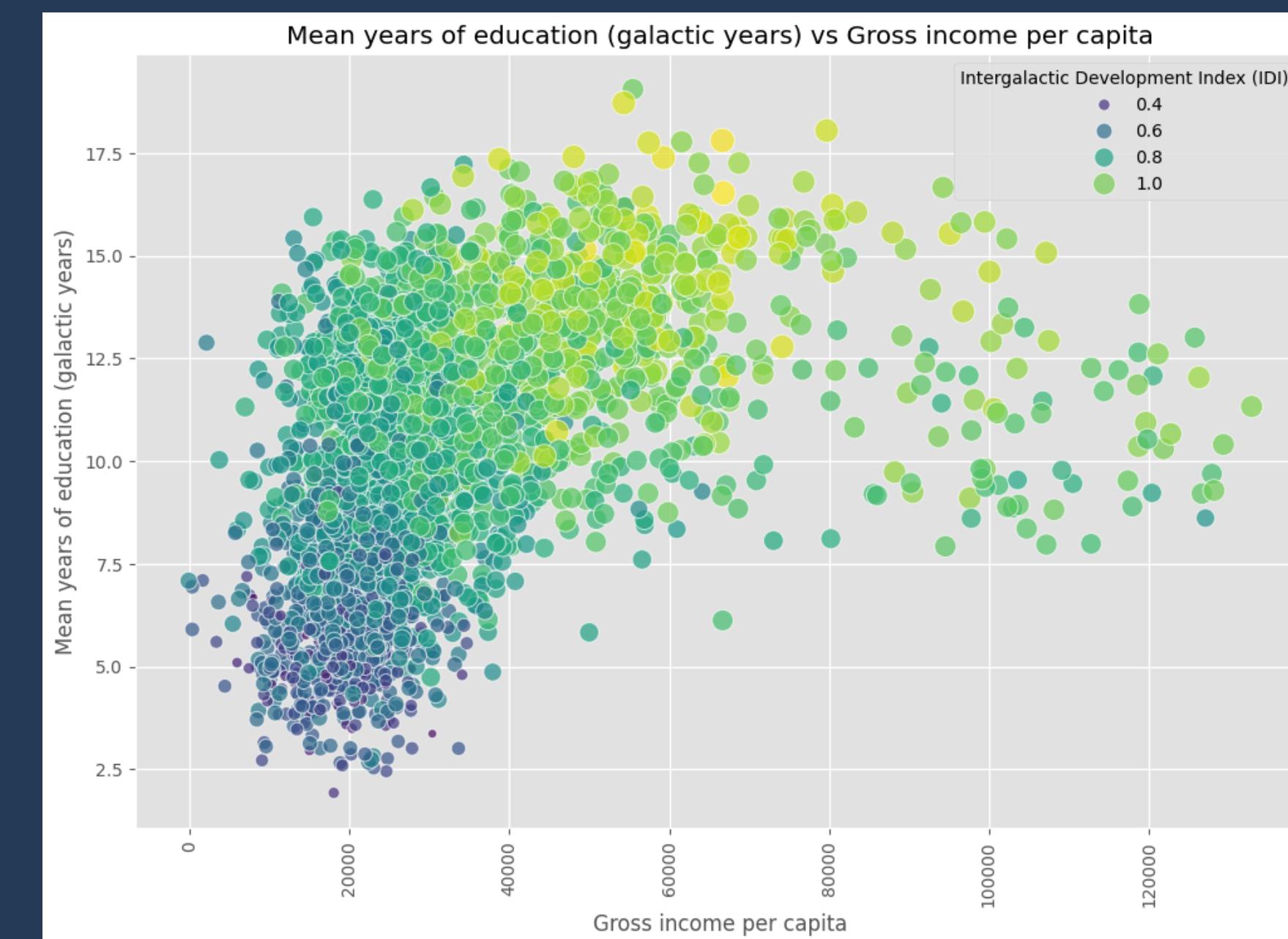
- The mean Well Being Index maintained a steady average around 0.08 until 1009020 where it started to increase, peaking at 1011030 at 0.127. The two last galactic years, while showing higher indices, also had the least observations.

THE INTERGALACTIC DEVELOPMENT INDEX

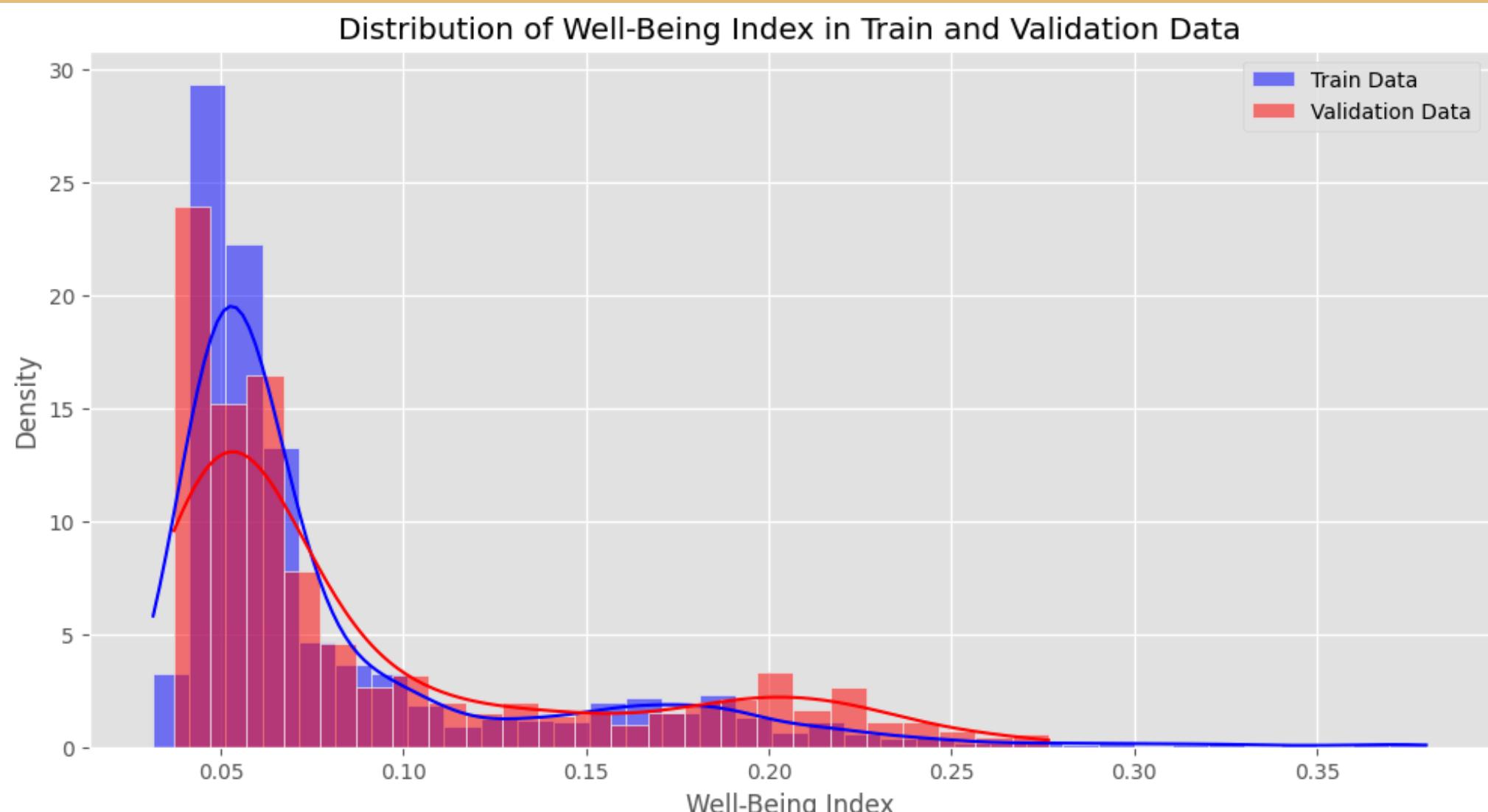


- There is a negative relationship between the Well-Being Index. This means the higher the Well-Being index eg 0.35, the lower the rank, eg 9.9
- The relationship between, however, is not fully observable from the plot as there are other factors that may be causing lots of the Well-Being Index to center around 0.05 regardless of the IDI rank.

- The galaxies with lower mean years of education, also had lower gross income per capita and Intergalactic Development Index
- Higher mean years of education (7.5 to 15) reported higher Intergalactic Development indices (0.8+) despite varying gross income per capita.



CONCLUSION AND INSIGHTS



- While many features were presented at the beginning, not all were useful in analysing the Well-Being Index of the individual observations. More complete data would have yielded more specific outcomes.
- The inverse relationship of the Intergalactic Development Index Rank with every other of the remaining features shows that increase in these features/indices would lead to a decrease in the figure of the rank but in literal sense they performed better, the lower the numerical rank.
- The 7 latest galactic years had fewer observations though higher mean Well-Being Index. It may be of concern to check whether this was an effect of mortality rates, expectancy rates or any other factors that may affect their reproducability
- The distribution of the well-being of validation data follows the same skewness pattern of the train data. Most of the observations of the well-being index of the validation data-set are also centered between 0.5 just like the train data.

