memo

|  |
| --- |
|  |

|  |  |
| --- | --- |
| To: | Professor Deborah Landowski |
| From: | Leonard Armstrong, David Ladd |
| Date: | 2021-Feb-13 |
| Re: | IST-622 Final Project Proposal |
|  | Proposal The team of Leonard Armstrong and David Ladd will investigate the world of one-sixth scale action figure collecting. American children born in the 1960s experienced the advent of the action figure with a toy named *G.I. Joe*®. The original G.I. Joe was a movable, posable figure slightly under 12-inches tall, making him one-sixth scale compared to the average male’s height. From Joe’s introduction in 1964 through today, dozens, if not hundreds, of manufacturers have produced one-sixth scale figures expanding far beyond G.I. Joe’s military themes. Data Sources Our investigation will use a combination of structured and unstructured data sources. From the structured data source domain, we will use ***Joebase***, a dataset taken from a relational database of a private collection of 500 one-sixth scale action figures. Joebase relates action figures to retail products[[1]](#footnote-1) and the price, year of release, manufacturer, the seller (retailer or gifter) of those products. Additionally, the data relates action figure products to genres or themes and to their product lines or series as defined by the manufacturer.  Unstructured information will come from ***eBay*** search result pages. The team will generate HTML results from specific eBay searches and use then scrape the HTML for relevant information. Possible Questions While the list of possible questions to be researched and reported is developing, some questions may logically follow from the results of Leonard Armstrong’s IST-622 Homework #1. The research related in that deliverable showed what appeared, to the naked eye, to be a superlinear rise in average figure prices over time. One follow-on research question is “*What function best describes the growth over that period and how well is it predicting current market prices?*” A second follow-on research topic may be the generation of a heatmap of manufacturers and/or genres and/or sellers over time to visualize more granular patterns in any of those three areas over the years.  On the unstructured data realm, eBay searches can be performed to review the current average sales for figures from higher-volume manufacturers, or for selected products in the Joebase collection. Also, since eBay lot titles are self-reported and tend to contain additional and non-essential verbiage[[2]](#footnote-2); research in fuzzy searching may be performed to match product names in Joebase to eBay lots. |

1. In most, but not all cases a retail product consists of one action figure. However, some products do contain more than one figure. [↑](#footnote-ref-1)
2. Added phrases like “Rare”, “Don’t miss out” and “No reserve!” come to mind. [↑](#footnote-ref-2)