“Simplified Chinese” abbreviated to SC, “Traditional Chinese” to TC

**Descriptive Statistics**

1. **Game Distribution & Publication Trends**

* Number of games published categorised by years
* Distribution of games by price range
  + Find the maximum price in the dataset for USD and CNY
  + For USD, start with <10 usd, adding 10 usd for each category
  + For CNY, start with <50 cny, adding 50 CNY for each category
  + Exclude game with ‘unknown price’ but set a category for free to play game
* Proportion of free-to-play vs. paid games over time

1. **Localisation Patterns**

* Number of games with any form of SC/TC localisation compared to English
* **Detailed localisation breakdown**: Percentage of games with:
  + No Chinese support
  + Interface/subtitles support
  + Full audio support
* **Localisation adoption trends**: number of new releases games with Chinese support by year.
  + Categorised by only has SC, only has TC and has both SC and TC
  + Adding these three categories together and compute ratio of Chinese localised game compared to total game released that year, supplement with ratio for each cateogory
* **Localisation by ‘indie’:** Compare Chinese localisation ratio between games with ‘indie’ tag in the *genres* variable and games without ‘indie’ tag.

1. **Review Metrics & Comparisons**

* **Review distribution**:
  + Proportion of total reviews that are in SC vs. English vs. other languages by years
  + Average positive review ratio in SC and English
* **Review volume comparison**: Average number of reviews per game in English vs. SC by years
* **Sentiment gap analysis**: Average difference between positive review ratios in SC vs. English for the same games by years

1. **Price Analysis**

* **Regional price comparison**: Average USD vs. CNY pricing by years
  + Check maximum price for USD and CNY in the dataset (done in **Game Distribution** section as well)
  + Check number of games with unknown price
* Price check in both USD and CNY.
  + **Price to localization relationship**: Do games with Chinese localisation command higher prices in general? Analysis by years.
  + **Price to indie relationship**: What is the average price for indie game every year? What is the difference between indie or non-indie?
  + **Price to Early Access relationship**: Does game with Early Access tag tends to have indie tag as well? What is the average price for them?
* **Regional price differences**: Do games increase in their average price every year? How this price differences in USD and CNY?

1. **User Tags & Categorization**

* Frequency for each user tag/genres
* **Genre-specific localisation**: Which user tag/genres have highest/lowest Chinese localisation rates
* **Tag association with localization**: Most common user tags/genres for games with/without SC or SC localisation
* **Genre preferences**: Top 10 User tag/genres with highest proportion of SC reviews
* **Cultural preference indicators**: Tags that show significant differences in positive review ratios between SC and English reviews

1. **~~Publisher/Developer Patterns~~**

* ~~Top developers/publishers by percentage of catalog with Chinese localisation~~
* ~~Top developers/publishers by absolute number of games with Chinese localisation~~
* ~~Comparison of Chinese vs. non-Chinese developers in terms of localization and review patterns (?)~~