Problem Set #1

MACS 30100, Dr. Evans
Chih-Yu Chiang

Problem 1

Part (a). Find an article

Game Changer: The Topology of Creativity

This article examines the sociological factors that explain why some creative teams are able to produce game-changing culture products. It applies empirical data collected from video game industry; video game like art, film, and dance is often deemed as an expression of culture.

Part (b). Detailed citation

De Vaan, M., Stark, D., Vedres, B. (2015). *Game Changer: The Topology of Creativity.* American Journal of Sociology, 120(4), 151. http://doi.org/10.1086/681213

Part (c). Statistical model

This article implements a couple of OLS Regressions as its method. It employs Distinctiveness, Critical Acclaim, and Game Changer as dependent variables; each in one regression model. The models share an array of sociological factors as independent variables. The equation is exhibited as follows, demonstrated with Distinctiveness as the independent variable

$$\begin{aligned} Distinctiveness = \beta_0 + \beta_1(FoldedDiversity) + \beta_2(CognitiveDiversity) + \beta_3(StructuralFolding) \\ + \beta_4(Constraint) + \beta_5(MeanGroupSize) + \beta_6(MeanGroupSize)^2 + \beta_7(NumberofGroups) \\ + \beta_8(NumberofMembers) + \beta_9(NumberofNewbies) + \beta_{10}(GamesTenure) + \beta_{11}(PastReviewScore) \\ + \beta_{12}(HighPerformers) + \beta_{13}(StarDeveloper) + \beta_{14}(SingleFirm) + \beta_{15}(MeanFirmAge) \\ + \beta_{16}(NumberofElements) + \epsilon \end{aligned}$$

Part (d). Endogenous and Exogenous variables

- Endogenous variables: $Distinctiveness$, $CriticalAcclaim$, and $GameChanger$.
- Exogenous variables: $FoldedDiversity$, $CognitiveDiversity$, $StructuralFolding$, $Constraint$, $MeanGroupSize$, $MeanGroupSize^2$, $NumberofGroups$, $NumberofMembers$, $NumberofNewbies$, $GamesTenure$, $PastReviewScore$, $HighPerformers$, $StarDeveloper$, $SingleFirm$, $MeanFirmAge$, and $NumberofElements$.

Part (e). Classify the model

- Static: This model describes a static relationship between the variables; theres no interactions like intertemporal of spatial ones between each observation.
- Linear: OLS regressions assume a linear dependent and independent variable relationship.

- Stochastic: OLS regressions assume an error term following $N(0, \sigma^2)$ in each model, which brings about stochasticity.

Part (f). Possible variable of feature

- $AverageMemberAge$: Young people are often considered having more creative power, with more passion, and willing to take more risk in their works; those characteristics could contribute to developing game-changing products.
- $AverageWorkingHourperMember$: One creative productivity source is exposing one selves to the various things in everyday life. Average working hour focusing on the focal project could be a reversal proxy to this.

Problem 2

Part (a)(b). Model for popular musicians live

This is a model estimating popular musicians lifespan.

$$\begin{aligned} PredictedLifespan = \beta_0 + \beta_1(NumberofRelationshipScandals) + \beta_2(AlcoholConsumption) \\ + \beta_3(DietPreference) + \beta_4(TravelTimeperActiveYear) + \beta_5(Spouse) \\ + \beta_6(Religion) + \beta_7(EducationLevel) + \beta_8(Nationality) + \beta_9(Sex) + \beta_{10}(YearBorn) + \epsilon \end{aligned}$$

Some definitions:

- $AlcoholConsumption$: 1 (least preference) 5 (strongest preference) scale. Might be acquired by asking the people close to the musician.
- $DietPreference$: 1 (least healthy) 5 (strongest healthy) scale. Might be evaluated by the musicians favorite foods.
- $TravelTimeperActiveYear$: The length of time not living in the musicians town of residency each year.
- $Spouse$: Binary. If the musician has a spouse at the time the model is estimated, 1; otherwise, 0.
- $Religion$: Binary. If the musician has a stable religious belief at the time the model is estimated, 1; otherwise, 0.

Part (c). Key factors

The variables illustrating the musicians life style is healthy or not, including $NumberofRelationshipScandals$, $AlcoholConsumption$, $DietPreference$, $TravelTimeperActiveYear$, $Spouse$, and $Religion$ are the key variables.

On top of that, some demographic factors can also influence life span, and therefore worth being included. They are $EducationLevel$, $Nationality$, Sex, and $YearBorn$.

Part (d). Reason of choosing factors

After randomly collecting 10 passed-away popular musicians information (refer to the table here), I found that many of them died young and because of the sickness or reasons resulting from bad life styles (not natural reasons). Out of the 10 musicians, 3 died directly and indirectly because of intoxication; 2 died because of heroin drugging.

It is therefore I included variables illustrating good or bad life styles. More $RelationshipScandals$, $AlcoholConsumption$, bad $DietPreference$, longer time traveling around could lead to high pressure, bad eating and resting, and thus shorter lifespan. On the other hand, having a $Spouse$, a good $Religion$, and knowledge from higher Education can lead to better everyday habits lead to longer life.

Part (e). Preliminary test

- Step 1. Collect a list of passed-away popular musicians from music services such as Spotify and online rankings such as this one for rock musicians.
- Step 2. Collect data of each musician; sort into database. Many of data of the variables employed in the model can be found on Wikipedia, such as Nationality, Education Level, and the passed-away musicians actual age when they died. For rest of the variables, such as Religion and Number of Relationship Scandals, could also possibly be found through google search, since our observing targets were all super stars and often left an abundance of personal life information in coverage.
- Step 3. Explore the data see if theres structural patterns between independent variables and passed-away musicians life span.
- Step 4. If the data size is sufficient, a regression model can be built according to it, using the actual age when the musician died as the dependent variable.
- Step 5. We can then examine the statistics of each parameter to see its significance, and the MSE (comparing predicted lifespan from this model and real age the musician died) to determine if the model as a whole provides a good estimation.