Module 1

Excel

* Conditional Format
  + Home – Conditional Formatting – Highlight cell rules – equal to – color
* Create new column
  + Adding new data
    - Formula = Round(E2/D2\*100,0)
    - Contional formatting – color scales – more rules – manage rules – 90th tile까지만 설정
* Formula <Round> 사용시 0으로 나눌수 없을시에 #Div/0! 에러 발생
  + 에러시, IFERROR formula 사용하여 Formula 2개를 동시에 적용가능, 에러 최소화. \*단, formula imput 순서가 중요(nesting formula라고 함)
* Pivot table
  + Insert – pivot table – 범위 선택 – pivot table fields에서 카테고리 설정
* Pivot chart
  + Pivot table click – 위에 pivot chart analyze 탭 선택 – pivot chart - design에서 차트 종류 및 외형 설정
  + Chart title : design - add chart element – chart title – 위치 선정
* Covert Unix timestamp to readable format
  + Formula : =(((J2/60)/60)/24)+DATE(1970,1,1)
    - (1970,1,1) = **epoch**
* Vlookup
  + 특정 데이터만을 원하여 표시하길 원하는 경우 Vlookup 포뮬러를 사용한다.
    - 먼저 새로운 sheet을 만들고 column에 맞는 타이틀을 넣는다.(ex. Name, blurb)
    - 그리고 Name 아래 표기하기 원하는 제목을 대소문자 띄어쓰기 정확하게 표기
    - 그리고 formula =VLOOKUP(A2, Kickstarter!B:C, 2, FALSE)
      * A2는 데이터를 찾기 바라는 타이틀을 가르키는것
      * Kickstarter!는 sheet의 이름 + !
      * B:C는 어느 column부터 어느 column까지 인가
      * 숫자 2는 표시하기 원하는 데이터 column이 왼쪽에서 몇번째인가(ex. Blurb를 표시하려면 왼쪽에서 2이기에 숫자 2.)
* Mean, Median, Mode
  + Mode = 가장 quantity가 많은 것
  + Mean 과 Median이 가까울 때, 데이터는 Symmetric(Half is above mean, half is below mean)
  + Mean 과 Median이 많이 다르다면, 데이터는 Skewed(extreme values are pulling the mean higher or lower)
  + If mean is much **higher** than median, it is **skewed to the right**. If mean is much **lower** than median, it is **skewed to the left**.
  + Mean formula : =Average(‘title of the data sheet’!D:D) D:D = which column
  + Median formula : = Median(‘title of the data sheet’!D:D)
* Standard Deviation, Quartiles, IQR(Interquartile Range)
  + Standard Deviation formula : =STDEV.P('Title of the data sheet'!D:D)
  + Quartiles
    - Lower quartile : 25%, Upper quartile : 75%
    - Formula : =QUARTILE.EXC('Title of the data sheet'!D:D, 3) the number by the end : 1=lower quartile, 3=upper quartile
  + IQR = Upper quartile – Lower Quartile
* Outliers
  + Definition : it is an extreme point of data where they are much larger than the rest of the data or much smaller.
  + Two ways to find out. 1) Mean & standard deviation 2) Median & interquartile range(IQR)
  + 1) using first method ; more than 3 standard deviations higher or lower than the mean is considered as an outlier
  + 2) using second method ;
    - any value greater than the upper quartile(75%) plus 1.5 x IQR is considered as an outlier
    - any value less than the lower quartile(25%) minus 1.5 x IQR is considered an outlier

**\*\* IQR method is preferred most of the time**