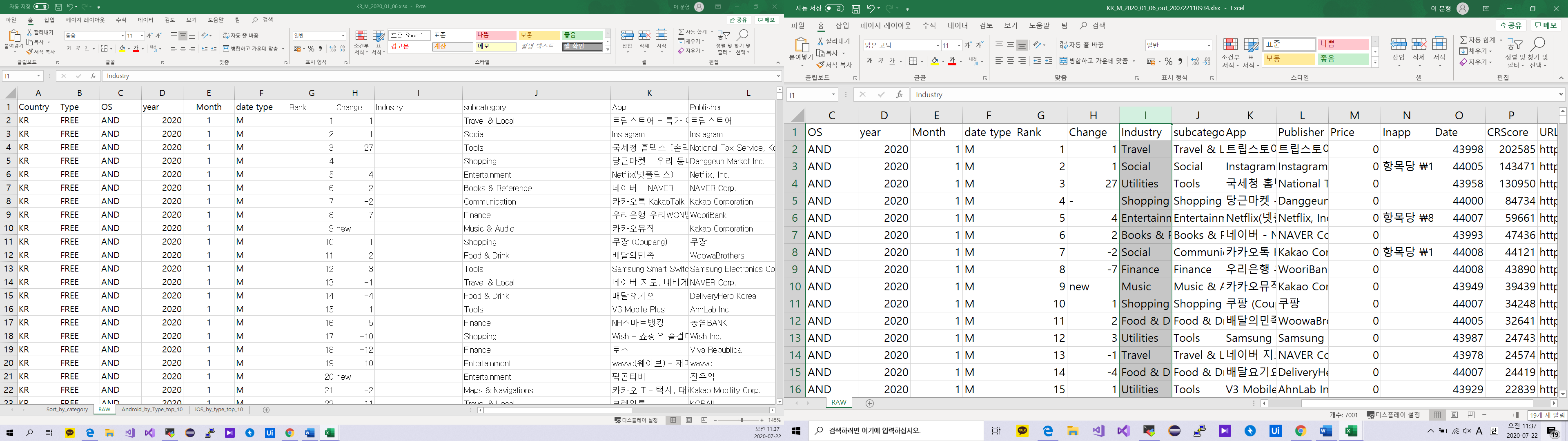
프로그램명 : check\_category

제작자 : 이문형

연락처 : [moonstar114@naver.com](mailto:moonstar114@naver.com)

1. 프로그램 기능 소개



후

전

프로그램을 실행하면, Industry 열을 \*Industry 카테고리 분류표에 기반하여 업데이트한다.

\* Industry 카테고리 분류표 3-1) 참고

2. 프로그램 사용법

2-1) 환경

윈도우 환경 : Window 10

python 3.5 ( using pandas, numpy, openpyxl )

2-2) 사용법

(1) 준비

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Country | Type | OS | year | Month | date type | Rank | Change | Industry | subcategory | App | Publisher | Price | Inapp | Date | CRScore | URL(Address) |

위와 같은 필드를 가진 액셀 파일이 필요합니다. 이는 캘커타 CMS (<http://203.245.4.81:8888/> )로부터 다운로드 받은 CSV 파일에서 \*combine\_file 프로그램을 통해 얻을 수 있는 파일입니다.

( \* combine\_file 프로그램 : CMS에서 받은 CSV 파일에 Country, Type, OS, year, Month, date type과 Industry 열을 추가해주는 프로그램이다. 이 프로그램은<https://github.com/m00nstarr/combine_data>에 공유 되어 있습니다.)



확인 항목 :

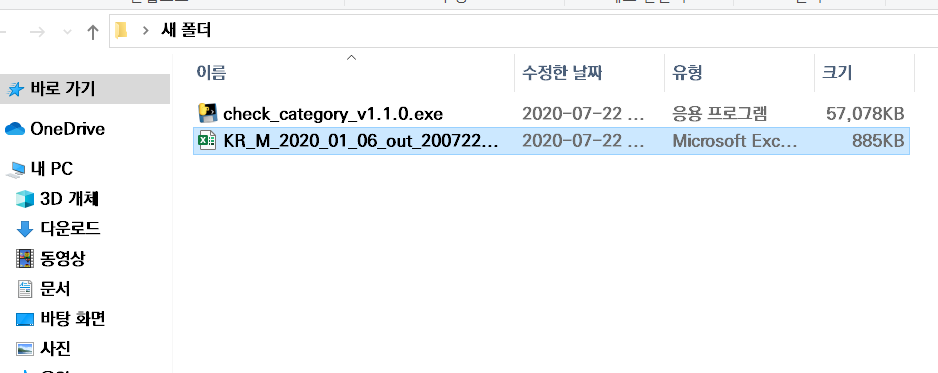
1. 위 필드를 가진 액셀 파일 (확장자 : xlsx )의 시트 이름이 ‘RAW’ 여야 합니다.

1 시트 이름 RAW 확인



2. Industry 열이 I열에 위치하여야 합니다.

3. Industry 열을 업데이트 할 파일을 exe 파일이 있는 폴더에 위치시킵니다.



2) 실행

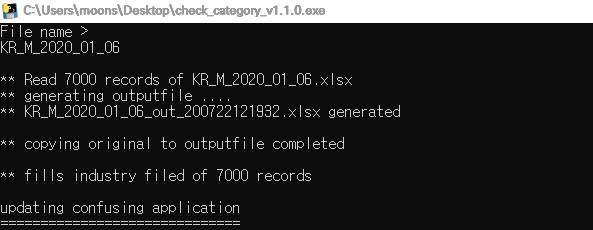
1. check\_category\_v.1.1.0.exe 파일을 더블 클릭하여 실행합니다.

다음과 같은 화면이 위치하면



업데이트 할 데이터 합본의 파일명을 입력합니다. 이 때, 확장자는 입력할 필요 없습니다.

2.

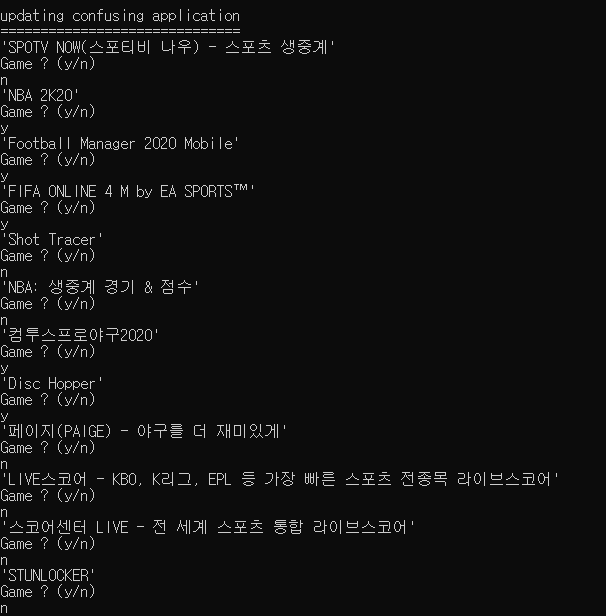


액셀 내의 레코드 개수를 출력하며 output 파일을 생성합니다.

output 파일의 이름 : (원본\_파일\_이름)\_out\_(연월일시분초).xlsx

생성한 이후엔, 원본 파일을 output 파일로 복사를 합니다. 복사가 완료되면, 각 record 별 subcategory를 통해 industry 필드를 업데이트 합니다.

3.

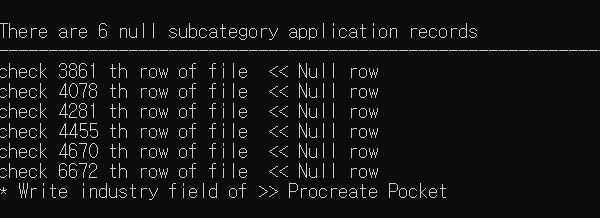


sports 서브카테고리는 Games와 sports가 혼재되어 있습니다. 따라서, 직접 분류를 해주어야 합니다. sports 서브카테고리의 row 별로 질의를 하게 되고 Game 어플리케이션이면 y, sports이면 n을 입력합니다.

4.

원본 데이터 내에서 subcategory가 비어있거나, 행 자체가 NULL값을 가지고 있을 수 있습니다.

이 과정에서는 비어있는 row들의 index를 알려주며, subcategory가 비어있는 경우에는 직접 입력해주어야 합니다.



위 사진의 예시에서는 Procreate Pocket의 subcategory가 비어있는 경우입니다. 이 상황에선 Procreate Pocket에 알맞은 industry 필드인 Productivity를 입력해주시면 됩니다. 이후엔 Procreate Pocket의 subcategory 필드를 직접 입력하고 아웃풋파일을 저장하며 프로그램이 자동종료 됩니다.

3. 참고

3-1) Industry 카테고리 분류표

|  |
| --- |
| **Books & Reference** |
| Books |
| Books & Reference |
| Reference |
| **Business** |
| Business |
| **Education** |
| Education |
| **Entertainment** |
| Comics |
| Entertainment |
| **Finance** |
| Finance |
| **Food & Drink** |
| Food & Drink |
| **Games** |
| Action |
| Adventure |
| Arcade |
| Board |
| Card |
| Casino |
| Casual |
| Games |
| Puzzle |
| Racing |
| Role Playing |
| Simulation |
| Sportsgame |
| Strategy |
| Trivia |
| Word |
| **Health** |
| Health & Fitness |
| Healthcare & Fitness |
| Medical |
| **Kids** |
| Educational |
| Kids |
| **Lifestyle** |
| Auto & Vehicles |
| Events |
| Family |
| House & Home |
| Lifestyle |
| Parenting |
| **Music** |
| Music |
| Music & Audio |
| **Navigation** |
| Maps & Navigations |
| Navigation |
| **News** |
| News |
| News & Magazines |
| **Photo & Video** |
| Photo & Video |
| Photography |
| Video Players & Editors |
| **Productivity** |
| Art & Design |
| Personalization |
| Productivity |
| **Shopping** |
| Shopping |
| **Social** |
| Communication |
| Dating |
| Social |
| Social Networking |
| **Sports** |
| Sports |
| **Travel** |
| Travel |
| Travel & Local |
| **Utilities** |
| Tools |
| Utilities |
| **Weather** |
| Weather |