```
In [1]: import pandas as pd
```

```
In [2]: file_path_titles = "../gitignore/title.akas.tsv"
        data_titles = pd.read_csv(file_path_titles, sep="\t")
```

```
In [3]: data_titles_copy = data_titles.copy()
```

```
In [4]: total_1 = data_titles.shape[0]
        print(f"Total: {total_1}")
        data_titles.nunique()
```

```
In [5]: data_titles_fr = data_titles[data_titles['region'] == 'FR']
```

```
In [6]: data_titles_autres = data_titles[~data_titles['titleId'].isin(data_titles_fr['ti
```

```
In [7]: data_titles_autres = data_titles_autres[data_titles_autres['isOriginalTitle']==1
```

```
In [8]: data_titles_final = pd.concat([data_titles_fr,data_titles_autres], axis=0)
```

```
In [9]: # Grouper par 'titleId' et consolider les titres dupliqués dans une liste
        consolidated_titles = data_titles_final.groupby('titleId')['title'].apply(list).
```

```
In [10]: # Fusionner les titres consolidés avec le dataframe d'origine, en éliminant les
         data_titles_final_unique = data_titles_final.drop_duplicates(subset='titleId')
         data_titles_final_unique = data_titles_final_unique.merge(consolidated_titles.re
```

```
In [11]: display(data_titles_final_unique [data_titles_final_unique['titleId'] == 'tt3775
```

	titleId	ordering	title	region	language	types	attributes	isOrigi
3778371	tt3775086	20	On l'appelle Jeeg Robot	FR	\N	imdbDisplay	\N	

```
In [12]: export = "../gitignore/title.akas_final.tsv"
         data_titles_final_unique.to_csv(export, sep="\t", index=False)
```