

Change log

Version 1.0.0

Between our initial upload of the data and the end of our grace period for editing the files, we fixed a small number of errors which were included within the 1.0.0 release.

Affected records	Action	Explanation
Persons P53365 and P23819 on film F88805	Delete	P53365 is a duplication of P33552, credited for the same role, due to a typo in the given name. P23819 is a duplication of P86278, credited for the same role, due to a missing dot in middle initial.
Person P28834 on film F65075	Amend	This person's credited-as name was misspelled so we corrected it.
Films F27590 and F57074.	Amend	The titles of these films did not follow the naming scheme used for film titles (they appeared as "THE FILM" rather than "FILM, THE") so we corrected them to match the scheme used throughout.
Film F24408	Amend	We changed the main title of this film to "UMBRELLA WOMAN, THE", as this seems to be the primary name by which the film is known, and we moved the previous value to the other titles field.

Version 1.0.1

In this patch, we made the following changes to the dataset:

Affected records	Action	Explanation
Person P21378 on film F62336	Delete	P21378 is a duplicate of P59130, credited for the same role, due to a typo in the given name. The source of P59130 was also modified to "sg" to reflect that the incorrect data came from the Screen Guide, which is higher in our source hierarchy than IMDb.
Person P32159 on film F10508	Delete	P32159 is a duplicate of P31578, credited for the same role, due to a typo in the given name.
Person P72641 on film F73329	Delete	P72641 is a duplicate of P63071, credited for the same role, due to a typo in the given name.
Person P59171 on film F29178	Delete	P59171 is a duplicate of P28694, credited for the same role, due to a typo in the given name. The source of P28694 was also modified to "sg" to reflect that the incorrect data came from the Screen Guide, which is higher in our source hierarchy than IMDb.
Person P82775 on film F49033	Delete	P82775 is a performing name for P70210, also credited for the same role, so this is essentially a duplicate.
Person P37706 on film F75222	Delete	P37706 is a performing name for P77802, also credited for the same role, so this is essentially a duplicate. The source of P77802 was also modified to "sg" to reflect that the incorrect data came from the Screen Guide, which

		is higher in our source hierarchy than IMDb.
Person P65454 on film F15602	Delete	P65454 is a performing name for P78394, also credited for the same role, so this is essentially a duplicate. The source of P78394 was also modified to "sg" to reflect that the incorrect data came from the Screen Guide, which is higher in our source hierarchy than IMDb.
Person P96949 on film F24071	Delete	P96949 is a company that is synonymous with P24980, also credited for the role, so this is essentially a duplicate.
Person P91800 on film F63794	Delete	This is a "based on the epic poem by" credit and thus outside of scope, it should have been removed during data cleaning.
Person P31264 on film F36566	Delete	P31264 is a company, not a person.
Person P53397 on film F78580	Delete	P53397 is a company, not a person.
Persons P25323 and P76043 on film F32448	Delete	These credits (which appear on IMDb as "Furber Meyer" and "Charig") appear to be references to the use of the song "That's a Good Girl" in the film, which is written by Furber, Meyer, and Charig.
Person P49305 on film F34871 segment F3487112	Delete	P49305 is a company, not a person.
Person P24014 on film F34149	Delete	P24014 is a company, not a person.
Person P83276 on film F54100	Amend	P83276 is the same person as P22544, though their names are spelled differently across credits. We changed the person ID from P83276 to P22544 to ensure these roles are grouped under the same identifier.
Person P43752 on film F30086	Amend	We don't think that the P43752 on film F30086 is the same person as the P43752 on film F78555 so we generated a new ID P79714 for the case on film F30086.
Persons P60841 and P55563 on film F48848	Amend	These credits both refer to "Rene and Rochford", though the name is spelled differently in each. Oz Movies has the head credits for this film which shows that "Rochford" is the correct spelling, but that the credit was for "Art direction and wardrobe design". We thus kept the row for the "Costume Designer" role P60841 and deleted the "Production Designer" role P55563 (we have otherwise treated art direction as a different role to production design) and set the spelling to "Rochford".
Persons P57341 and P19868 on film F21897	Amend	For some reason Brad Lindenmayer's editing credit for F21897 appears on IMDb as two credits "Brad Schmad" (P57341) and "Lidenmayer" (P19868), and we missed this during data cleaning. We deleted the role with ID P57341 and changed the ID of the remaining role from P19868 to P17150, which is the person's ID elsewhere in the dataset.
Person P53690 on film F34518	Amend	The credited-as name for P53690 was incorrect so we fixed it to match the credits.

We also made some changes to the way that bands and musical groups are captured in the data. Currently, we have two cases where we know that the "music by" credit was given to a band, but they appear in the data as individual band members. However, this is not consistently applied to the rest of the bands in the data. Rather than split all bands into their individual members, which is a difficult task (information is uneven for many of these bands) and which would add noise to the data, we instead decided to recombine the two bands into a single credit. Rather than remove bands from the data altogether (we think these credits can also reveal interesting aspects of collaboration in the industry) we decided to modify the IDs of these bands to end with the suffix "(G)" so that they are easy to identify and filter if desired. The affected records in the data are as follows:

Affected records	Action	Explanation
Persons P25247, P37226 and P80881 on film F64208	Delete	These three credits for band members were deleted to reduce the credit to a single role.
Person P19032 on film F64208	Amend	This credit is re-used as the credit for the band as a whole, so its ID is changed to P19032(G).
Persons P81977 and P28660 on film F26314	Delete	These two credits for band members were deleted to reduce the credit to a single role.
Person P33457 on film F26314	Amend	This credit is re-used as the credit for the band as a whole, so its ID is changed to P33457(G).
Persons P33457, P56628, P53690, P39668, P51696, P73075, P79410, P76527, P94503, P68256, P46492, P69348, P55428, P46890, P97294, P43045, P12744, P83774, P89007, P12851 and P68711.	Amend	These IDs all identify bands and musical groups, so they were changed to add the "(G)" suffix to the end of the IDs.

Finally, in this release we have updated the technical documentation to reflect these changes and provide more details on the data versioning process for the dataset.