Data exploration and cleaning

This notebook corrects the errors detected by <1.1_quality_check.Rmd>, based on the feed back from the people who collected the data. Corrections are done within this script to ensure reproducibility.

The output is a clean kobo file that can be used for analyses.

## Get data, libraries and functions

Load required libraries:

library(tidyr)  
library(dplyr)  
library(utile.tools)  
library(stringr)  
library(ggplot2)

Useful functions

# not in  
"%!in%" <- function(x, y)!('%in%'(x,y))

Get Kobo raw output data:

kobo\_output<-read.csv(file="International\_Genetic\_Indicator\_testing\_V\_4.0\_-\_latest\_version\_-\_False\_-\_2023-06-23-14-59-49.csv", sep=";", header=TRUE) %>%  
  
## add taxon column  
mutate(taxon=(utile.tools::paste(genus, species, subspecies\_variety, na.rm=TRUE))) %>%  
 # remove white space at the end of the name  
 mutate(taxon=str\_trim(taxon, "right"))

## Filter tests and records marked as not approved

Filter out records which were marked as “not\_approved” in the manual Kobo validation interface (this means country assessors determined the is something wrong with that particular record).

# check if any species is flagged as "validation\_status\_not\_approved"  
kobo\_output %>%  
 filter(X\_validation\_status=="validation\_status\_not\_approved")

## start end  
## 1 2022-11-27T11:08:34.622+02:00 2022-11-29T14:16:12.307+02:00  
## 2 2023-03-05T23:09:12.580-06:00 2023-03-06T01:38:00.206-06:00  
## 3 2023-03-01T14:50:06.988-06:00 2023-03-06T16:11:09.351-06:00  
## 4 2023-03-21T15:37:44.241-06:00 2023-03-21T17:33:36.631-06:00  
## 5 2023-04-25T14:20:22.049-06:00 2023-04-25T14:39:14.678-06:00  
## username email\_assessor name\_assessor  
## 1 username not found paulette.bloomer@up.ac.za Paulette Bloomer  
## 2 username not found rominamm@ciencias.unam.mx Romina Miranda Méndez  
## 3 username not found maritzga@ciwncias.unam.mx Ana Weiger  
## 4 username not found rominamm@ciencias.unam.mx Romina Miranda Méndez  
## 5 username not found ecn.robertson@colostate.edu Erica Robertson  
## country\_assessment genus species subspecies\_variety  
## 1 south\_africa Panthera leo   
## 2 mexico Pseudoeurycea Lineola   
## 3 mexico Ambystoma Altamirani 0  
## 4 mexico Castilleja Tolucensis   
## 5 united\_states Cambarus cracens   
## scientific\_authority common\_name taxonomic\_group  
## 1 (Linneaus, 1758) Lion mammal  
## 2 Cope, 1865 Tlaconete cola larga bryophyte  
## 3 Dugès, 1895 Ajolote arroyero de montaña amphibian  
## 4 Kunth, 1818 Garañona de alta montaña angiosperm  
## 5 R.W.Bouchard & Hobbs, 1976 Slenderclaw Crayfish invertebrate  
## GBIF\_taxonID NCBI\_taxonID national\_taxonID source\_national\_taxonID  
## 1 5219404 9689   
## 2 2431406 1708625 12873ANFIB Enciclovida  
## 3 2431955 354123 12770ANFIB SNBI  
## 4 3734635 ANGIOSPERMA Enciclovida  
## 5 2227363 1240897 ECOS: 9792   
## section3\_note n\_extint\_populations extint\_pop\_names n\_extant\_populations  
## 1 NA -999 5  
## 2 NA 2 Oaxaca, Zoquitlán 27  
## 3 NA 0 0  
## 4 NA 0 137  
## 5 NA -999 2  
## other\_populations why\_other  
## 1   
## 2 no\_more\_pops   
## 3 no\_more\_pops   
## 4 no\_more\_pops   
## 5 no\_more\_pops   
## time\_populations  
## 1 2008-2013  
## 2 1865-2021  
## 3 2019  
## 4 Al pertenecer a la flora de la FVTM, es una especie muy vulnerable a los cambios en el ecosistema producidos por el cambio climático y según el estudio de Ramírez-Amezcua et al. (2016), no se prevé una existencia en las partes altas de las montañas debido al cambio de temperatura en las regiones que habita. Además de la actividad humana que su paso por la montaña fragmenta su ecosistema (Arredondo-Amezcua, s.f.)  
## 5 1976-2017  
## defined\_populations  
## 1 management\_units  
## 2 other  
## 3 geographic\_boundaries adaptive\_traits  
## 4 genetic\_clusters  
## 5 eco\_biogeo\_proxies  
## source\_definition\_populations  
## 1 Subpopulations according to the Regional Red List Assessment (2016) and the Biodiversity Management Plan (2015). Only two large historical subpopulations (Kruger National Park and Kgalagadi Transfrontier National Park), not completely fenced and unmanaged. One large, older reintroduced population (Hluhluwe-iMfolozi Park), private reserves open to the Kruger National Park, a small naturally recolonized transfrontier park (Mapungubwe) and a managed metapopulation comprising 45 small fenced reserves/private properties.  
## 2 Hace falta información al respecto  
## 3 Dugues 1896 describes Abytsoma altamirani as a new species of mountain axolotl, based on the persistence of vomerian teeth in the row in adults, since in other species they were only present in larvae. Another characteristic is the back of the head which is modified by the persistent parts of the hyoid. Taking the above characteristics into account, Dunn (1928) considers A. altamirani to be different and should be placed in a genus that he describes as Rhyacosiredon, a criterion that is later supported by Taylor (1938), (Matías, 2006).  
## 4 Por capa geográfica, se contaron por montaña  
## 5 Populations delineated by Hydrological Unit Code (HUC) 12.  
## map\_populations  
## 1 Lion populations of South Africa-17\_25\_10.pdf  
## 2   
## 3 Ambystoma altamirani-15\_20\_49.png  
## 4 Castilleja\_tolucensis\_mapas-17\_17\_10.jpg  
## 5 Screen Shot 2023-04-25 at 2.32.59 PM-14\_33\_31.png  
## map\_populations\_URL  
## 1 https://kc.kobotoolbox.org/media/original?media\_file=amastretta%2Fattachments%2Fb777f0bf79924d45ab7157705a420e5d%2F01aa2ec2-f7e0-4966-900b-1c4cf9e739cc%2FLion\_populations\_of\_South\_Africa-17\_25\_10.pdf  
## 2   
## 3 https://kc.kobotoolbox.org/media/original?media\_file=amastretta%2Fattachments%2Fb777f0bf79924d45ab7157705a420e5d%2F0d042a15-9c72-415a-b088-55c1a8a9f18b%2FAmbystoma\_altamirani-15\_20\_49.png  
## 4 https://kc.kobotoolbox.org/media/original?media\_file=amastretta%2Fattachments%2Fb777f0bf79924d45ab7157705a420e5d%2Ff41dcf91-0adc-4c00-a6d7-986c44bea2c5%2FCastilleja\_tolucensis\_mapas-17\_17\_10.jpg  
## 5 https://kc.kobotoolbox.org/media/original?media\_file=amastretta%2Fattachments%2Fb777f0bf79924d45ab7157705a420e5d%2F37576a4b-38da-4d28-941e-f6ec75c86741%2FScreen\_Shot\_2023-04-25\_at\_2.32.59\_PM-14\_33\_31.png  
## habitat\_decline\_area  
## 1 Lion used to occur throughout most of South Africa. By the 1900's only two populations remained. Decline of 80-97%.  
## 2 2,200 km2  
## 3 The population in general is not known, however, the species Ambystoma altamirani is found in an altitudinal range of 2,700 to 3,600 masl (Uribe et al., 2000).  
## 4   
## 5   
## source\_populations  
## 1 Bauer H, Chapron G, Nowell K, Henschel P, Funston P, Hunter LTB, Macdonald D, Packer C. 2015. Lion (Panthera Leo) populations are declining rapidly across Africa, except in Intensively Managed Areas. Proceedings of the National Academy of Sciences of the United States of America, 112(48), 14894–14899; \nFunston PJ, Levendal M. 2015. Biodiversity Management Plan for the lion (Panthera leo) in South Africa. Government Gazette No 39468; \nMiller S, Riggio J, Funston P, Power RJ, Williams V, Child MF. 2016. A conservation assessment of Panthera leo. In Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT, editors. The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa; \nRiggio J, Jacobson A, Dollar L, Bauer H, Becker M, Dickman A, et al. 2013. The size of savannah Africa: a lion's (Panthera leo) view. Biodiversity Conservation 22: 17–35.  
## 2 https://doi.org/10.22201/fc.25942158e.2022.1.394  
## 3 Lemos-Espinal, J. A., Smith, G. R., Ruíz, Á. H., & Ayala, R. M. (2016). Stream use and population characteristics of the endangered salamander, Ambystoma altamirani, from the Arroyo Los Axolotes, State of Mexico, Mexico. The Southwestern Naturalist, 61(1), 28-32.\nWooLRICH-PIñA, G., Smith, G. R., Lemos-Espinal, J. A., Zamora, A. E., & Ayala, R. M. (2017). Observed localities for three endangered, endemic Mexican ambystomatids (Ambystoma altamirani, A. leorae, and A. rivulare) from central Mexico. Herpetological Bulletin, 139, 12-15.\nHeredia-Bobadilla, R. L., Monroy-Vilchis, O., Zarco-González, M. M., Martínez-Gómez, D., Mendoza-Martínez, G. D., & Sunny, A. (2017). Genetic variability and structure of an isolated population of Ambystoma altamirani, a mole salamander that lives in the mountains of one of the largest urban areas in the world. Journal of Genetics, 96, 873-883.\n404 Not Found. (s. f.-b). https://iefectividad.conanp.gob.mx/i-efectividad/CyEN/PN+Desierto+de+los+Leones/Contexto+y+Planeaci%C3%B3n/Investigaci%C3%B3n+y+Monitoreo/Estudio+Ambystoma+altamirani.pdf  
## 4 https://enciclovida.mx/especies/153674-castilleja-tolucensis;  
## 5 U.S. Fish and Wildlife Service. (2019) Species Status Assessment Report for the Slenderclaw Crayfish (Cambarus cracens).  
## popsize\_data ne\_pops\_exists nc\_pops\_exists ratio\_exists  
## 1 yes other\_genetic\_info yes no  
## 2 data\_for\_species   
## 3 insuff\_data\_species   
## 4 insuff\_data\_species   
## 5 yes no\_genetic\_data yes no  
## species\_related ratio\_species\_related ratio\_year source\_popsize\_ratios  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## species\_comments  
## 1 In non-social species of the genus Panthera the Ne/Nc ratio is around 0.4.  
## 2   
## 3 Ne, Nc number is not specified, only that a population has a small size, which indicates the precarious conditions of its habitat  
## 4   
## 5   
## section5\_note kobo\_tabular pop\_tabular\_file pop\_tabular\_file\_URL  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## Name\_pop1 Origin\_pop1 IntroductionYear\_pop1 Ne\_pop1 NeLower\_pop1  
## 1 Kruger National Park natural NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 Short creek natural NA NA  
## NeUpper\_pop1 NeYear\_pop1 GeneticMarkers\_pop1 GeneticMarkersOther\_pop1  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## MethodNe\_pop1 SourceNe\_pop1 NcType\_pop1 NcYear\_pop1 NcMethod\_pop1  
## 1 Nc\_point 2015 Nc\_method\_count  
## 2   
## 3   
## 4   
## 5 Nc\_point 1970-1974 Nc\_method\_count  
## NcRange\_pop1 NcRangeDetails\_pop1 NcPoint\_pop1 NcLower\_pop1 NcUpper\_pop1  
## 1 1803 1715 1891  
## 2 NA NA NA  
## 3 NA NA NA  
## 4 NA NA NA  
## 5 90 NA NA  
## SourceNc\_pop1  
## 1 Ferreira SM, Funston PJ. 2010. Estimating lion population variables: prey and disease effects in Kruger National Park, South Africa. Wildlife Research 37:194–206; \nFerreira SM, Govender D, Herbst M. 2013. Conservation implications of Kalahari lion population dynamics. African Journal of Ecology 51:176–179;\nMiller SM, et al. 2013. Management of reintroduced lions in small, fenced reserves in South Africa: an assessment and guidelines. South African Journal of Wildlife Research 43:138–154;\nMiller S, Riggio J, Funston P, Power RJ, Williams V, Child MF. 2016. A conservation assessment of Panthera leo. In Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT, editors. The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa.  
## 2   
## 3   
## 4   
## 5 U.S. Fish and Wildlife Service. (2019) Species Status Assessment Report for the Slenderclaw Crayfish (Cambarus cracens).  
## Comments\_pop1  
## 1 Population 1: Kruger National Park only. System also part of Great Limpopo Transfrontier Park in east and private reserves to southwest (possibly adding another 500 individuals to the population). Population 2: South African part of Transfrontier Park only. Population 3: Population reintroduced in 1950s and augmented in late 1990s. Population 4: South African part of Transfrontier Conservation. Population 5: South African part of TFCA recolonised. Population 6: 45 small reserves reintroductions since 1990s.  
## 2   
## 3   
## 4   
## 5   
## Name\_pop2 Origin\_pop2 IntroductionYear\_pop2 Ne\_pop2  
## 1 Kgalagadi Transfrontier Park natural NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 Town creek natural NA  
## NeLower\_pop2 NeUpper\_pop2 NeYear\_pop2 GeneticMarkers\_pop2  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## GeneticMarkersOther\_pop2 MethodNe\_pop2 SourceNe\_pop2 NcType\_pop2 NcYear\_pop2  
## 1 NA Nc\_point 2015  
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA Nc\_point 1970-1974  
## NcMethod\_pop2 NcRange\_pop2 NcRangeDetails\_pop2 NcPoint\_pop2 NcLower\_pop2  
## 1 Nc\_method\_count 246 238  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 Nc\_method\_count 1 NA  
## NcUpper\_pop2  
## 1 256  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## SourceNc\_pop2  
## 1 Ferreira SM, Funston PJ. 2010. Estimating lion population variables: prey and disease effects in Kruger National Park, South Africa. Wildlife Research 37:194–206; \nFerreira SM, Govender D, Herbst M. 2013. Conservation implications of Kalahari lion population dynamics. African Journal of Ecology 51:176–179;\nMiller SM, et al. 2013. Management of reintroduced lions in small, fenced reserves in South Africa: an assessment and guidelines. South African Journal of Wildlife Research 43:138–154;\nMiller S, Riggio J, Funston P, Power RJ, Williams V, Child MF. 2016. A conservation assessment of Panthera leo. In Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT, editors. The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa.  
## 2   
## 3   
## 4   
## 5 U.S. Fish and Wildlife Service. (2019) Species Status Assessment Report for the Slenderclaw Crayfish (Cambarus cracens).  
## Comments\_pop2  
## 1 Population 1: Kruger National Park only. System also part of Great Limpopo Transfrontier Park in east and private reserves to southwest (possibly adding another 500 individuals to the population). Population 2: South African part of Transfrontier Park only. Population 3: Population reintroduced in 1950s and augmented in late 1990s. Population 4: South African part of Transfrontier Conservation. Population 5: South African part of TFCA recolonised. Population 6: 45 small reserves reintroductions since 1990s.  
## 2   
## 3   
## 4   
## 5   
## Name\_pop3 Origin\_pop3 IntroductionYear\_pop3 Ne\_pop3 NeLower\_pop3  
## 1 Hluhluwe-iMfolozi Park natural NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NeUpper\_pop3 NeYear\_pop3 GeneticMarkers\_pop3 GeneticMarkersOther\_pop3  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## MethodNe\_pop3 SourceNe\_pop3 NcType\_pop3 NcYear\_pop3 NcMethod\_pop3  
## 1 Nc\_point 2015 Nc\_method\_count  
## 2   
## 3   
## 4   
## 5   
## NcRange\_pop3 NcRangeDetails\_pop3 NcPoint\_pop3 NcLower\_pop3 NcUpper\_pop3  
## 1 120 NA NA  
## 2 NA NA NA  
## 3 NA NA NA  
## 4 NA NA NA  
## 5 NA NA NA  
## SourceNc\_pop3  
## 1 Ferreira SM, Funston PJ. 2010. Estimating lion population variables: prey and disease effects in Kruger National Park, South Africa. Wildlife Research 37:194–206; \nFerreira SM, Govender D, Herbst M. 2013. Conservation implications of Kalahari lion population dynamics. African Journal of Ecology 51:176–179;\nMiller SM, et al. 2013. Management of reintroduced lions in small, fenced reserves in South Africa: an assessment and guidelines. South African Journal of Wildlife Research 43:138–154;\nMiller S, Riggio J, Funston P, Power RJ, Williams V, Child MF. 2016. A conservation assessment of Panthera leo. In Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT, editors. The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa.  
## 2   
## 3   
## 4   
## 5   
## Comments\_pop3  
## 1 Population 1: Kruger National Park only. System also part of Great Limpopo Transfrontier Park in east and private reserves to southwest (possibly adding another 500 individuals to the population). Population 2: South African part of Transfrontier Park only. Population 3: Population reintroduced in 1950s and augmented in late 1990s. Population 4: South African part of Transfrontier Conservation. Population 5: South African part of TFCA recolonised. Population 6: 45 small reserves reintroductions since 1990s.  
## 2   
## 3   
## 4   
## 5   
## Name\_pop4 Origin\_pop4 IntroductionYear\_pop4 Ne\_pop4  
## 1 Mapungubwe National Park natural NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NeLower\_pop4 NeUpper\_pop4 NeYear\_pop4 GeneticMarkers\_pop4  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## GeneticMarkersOther\_pop4 MethodNe\_pop4 SourceNe\_pop4 NcType\_pop4 NcYear\_pop4  
## 1 NA Nc\_point 2015  
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcMethod\_pop4 NcRange\_pop4 NcRangeDetails\_pop4 NcPoint\_pop4 NcLower\_pop4  
## 1 Nc\_method\_count 10 NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NcUpper\_pop4  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## SourceNc\_pop4  
## 1 Ferreira SM, Funston PJ. 2010. Estimating lion population variables: prey and disease effects in Kruger National Park, South Africa. Wildlife Research 37:194–206; \nFerreira SM, Govender D, Herbst M. 2013. Conservation implications of Kalahari lion population dynamics. African Journal of Ecology 51:176–179;\nMiller SM, et al. 2013. Management of reintroduced lions in small, fenced reserves in South Africa: an assessment and guidelines. South African Journal of Wildlife Research 43:138–154;\nMiller S, Riggio J, Funston P, Power RJ, Williams V, Child MF. 2016. A conservation assessment of Panthera leo. In Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT, editors. The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa.  
## 2   
## 3   
## 4   
## 5   
## Comments\_pop4  
## 1 Population 1: Kruger National Park only. System also part of Great Limpopo Transfrontier Park in east and private reserves to southwest (possibly adding another 500 individuals to the population). Population 2: South African part of Transfrontier Park only. Population 3: Population reintroduced in 1950s and augmented in late 1990s. Population 4: South African part of Transfrontier Conservation. Population 5: South African part of TFCA recolonised. Population 6: 45 small reserves reintroductions since 1990s.  
## 2   
## 3   
## 4   
## 5   
## Name\_pop5 Origin\_pop5 IntroductionYear\_pop5 Ne\_pop5 NeLower\_pop5  
## 1 Small reserves natural NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NeUpper\_pop5 NeYear\_pop5 GeneticMarkers\_pop5 GeneticMarkersOther\_pop5  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## MethodNe\_pop5 SourceNe\_pop5 NcType\_pop5 NcYear\_pop5 NcMethod\_pop5  
## 1 Nc\_point 2015 Nc\_method\_count  
## 2   
## 3   
## 4   
## 5   
## NcRange\_pop5 NcRangeDetails\_pop5 NcPoint\_pop5 NcLower\_pop5 NcUpper\_pop5  
## 1 500 NA NA  
## 2 NA NA NA  
## 3 NA NA NA  
## 4 NA NA NA  
## 5 NA NA NA  
## SourceNc\_pop5  
## 1 Ferreira SM, Funston PJ. 2010. Estimating lion population variables: prey and disease effects in Kruger National Park, South Africa. Wildlife Research 37:194–206; \nFerreira SM, Govender D, Herbst M. 2013. Conservation implications of Kalahari lion population dynamics. African Journal of Ecology 51:176–179;\nMiller SM, et al. 2013. Management of reintroduced lions in small, fenced reserves in South Africa: an assessment and guidelines. South African Journal of Wildlife Research 43:138–154;\nMiller S, Riggio J, Funston P, Power RJ, Williams V, Child MF. 2016. A conservation assessment of Panthera leo. In Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT, editors. The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa.  
## 2   
## 3   
## 4   
## 5   
## Comments\_pop5  
## 1 Population 1: Kruger National Park only. System also part of Great Limpopo Transfrontier Park in east and private reserves to southwest (possibly adding another 500 individuals to the population). Population 2: South African part of Transfrontier Park only. Population 3: Population reintroduced in 1950s and augmented in late 1990s. Population 4: South African part of Transfrontier Conservation. Population 5: South African part of TFCA recolonised. Population 6: 45 small reserves reintroductions since 1990s.  
## 2   
## 3   
## 4   
## 5   
## Name\_pop6 Origin\_pop6 IntroductionYear\_pop6 Ne\_pop6 NeLower\_pop6 NeUpper\_pop6  
## 1 NA NA NA  
## 2 NA NA NA  
## 3 NA NA NA  
## 4 NA NA NA  
## 5 NA NA NA  
## NeYear\_pop6 GeneticMarkers\_pop6 GeneticMarkersOther\_pop6 MethodNe\_pop6  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## SourceNe\_pop6 NcType\_pop6 NcYear\_pop6 NcMethod\_pop6 NcRange\_pop6  
## 1   
## 2   
## 3   
## 4   
## 5   
## NcRangeDetails\_pop6 NcPoint\_pop6 NcLower\_pop6 NcUpper\_pop6 SourceNc\_pop6  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop6 Name\_pop7 Origin\_pop7 IntroductionYear\_pop7 Ne\_pop7  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NeLower\_pop7 NeUpper\_pop7 NeYear\_pop7 GeneticMarkers\_pop7  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## GeneticMarkersOther\_pop7 MethodNe\_pop7 SourceNe\_pop7 NcType\_pop7 NcYear\_pop7  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcMethod\_pop7 NcRange\_pop7 NcRangeDetails\_pop7 NcPoint\_pop7 NcLower\_pop7  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NcUpper\_pop7 SourceNc\_pop7 Comments\_pop7 Name\_pop8 Origin\_pop8  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## IntroductionYear\_pop8 Ne\_pop8 NeLower\_pop8 NeUpper\_pop8 NeYear\_pop8  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## GeneticMarkers\_pop8 GeneticMarkersOther\_pop8 MethodNe\_pop8 SourceNe\_pop8  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcType\_pop8 NcYear\_pop8 NcMethod\_pop8 NcRange\_pop8 NcRangeDetails\_pop8  
## 1   
## 2   
## 3   
## 4   
## 5   
## NcPoint\_pop8 NcLower\_pop8 NcUpper\_pop8 SourceNc\_pop8 Comments\_pop8 Name\_pop9  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Origin\_pop9 IntroductionYear\_pop9 Ne\_pop9 NeLower\_pop9 NeUpper\_pop9  
## 1 NA NA NA  
## 2 NA NA NA  
## 3 NA NA NA  
## 4 NA NA NA  
## 5 NA NA NA  
## NeYear\_pop9 GeneticMarkers\_pop9 GeneticMarkersOther\_pop9 MethodNe\_pop9  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## SourceNe\_pop9 NcType\_pop9 NcYear\_pop9 NcMethod\_pop9 NcRange\_pop9  
## 1   
## 2   
## 3   
## 4   
## 5   
## NcRangeDetails\_pop9 NcPoint\_pop9 NcLower\_pop9 NcUpper\_pop9 SourceNc\_pop9  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop9 Name\_pop10 Origin\_pop10 IntroductionYear\_pop10 Ne\_pop10  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NeLower\_pop10 NeUpper\_pop10 NeYear\_pop10 GeneticMarkers\_pop10  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## GeneticMarkersOther\_pop10 MethodNe\_pop10 SourceNe\_pop10 NcType\_pop10  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcYear\_pop10 NcMethod\_pop10 NcRange\_pop10 NcRangeDetails\_pop10 NcPoint\_pop10  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NcLower\_pop10 NcUpper\_pop10 SourceNc\_pop10 Comments\_pop10 Name\_pop11  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## Origin\_pop11 IntroductionYear\_pop11 Ne\_pop11 NeLower\_pop11 NeUpper\_pop11  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## NeYear\_pop11 GeneticMarkers\_pop11 GeneticMarkersOther\_pop11 MethodNe\_pop11  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## SourceNe\_pop11 NcType\_pop11 NcYear\_pop11 NcMethod\_pop11 NcRange\_pop11  
## 1   
## 2   
## 3   
## 4   
## 5   
## NcRangeDetails\_pop11 NcPoint\_pop11 NcLower\_pop11 NcUpper\_pop11 SourceNc\_pop11  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop11 Name\_pop12 Origin\_pop12 IntroductionYear\_pop12 Ne\_pop12  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NeLower\_pop12 NeUpper\_pop12 NeYear\_pop12 GeneticMarkers\_pop12  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## GeneticMarkersOther\_pop12 MethodNe\_pop12 SourceNe\_pop12 NcType\_pop12  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcYear\_pop12 NcMethod\_pop12 NcRange\_pop12 NcRangeDetails\_pop12 NcPoint\_pop12  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NcLower\_pop12 NcUpper\_pop12 SourceNc\_pop12 Comments\_pop12 Name\_pop13  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## Origin\_pop13 IntroductionYear\_pop13 Ne\_pop13 NeLower\_pop13 NeUpper\_pop13  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## NeYear\_pop13 GeneticMarkers\_pop13 GeneticMarkersOther\_pop13 MethodNe\_pop13  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## SourceNe\_pop13 NcType\_pop13 NcYear\_pop13 NcMethod\_pop13 NcRange\_pop13  
## 1   
## 2   
## 3   
## 4   
## 5   
## NcRangeDetails\_pop13 NcPoint\_pop13 NcLower\_pop13 NcUpper\_pop13 SourceNc\_pop13  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop13 Name\_pop14 Origin\_pop14 IntroductionYear\_pop14 Ne\_pop14  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NeLower\_pop14 NeUpper\_pop14 NeYear\_pop14 GeneticMarkers\_pop14  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## GeneticMarkersOther\_pop14 MethodNe\_pop14 SourceNe\_pop14 NcType\_pop14  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcYear\_pop14 NcMethod\_pop14 NcRange\_pop14 NcRangeDetails\_pop14 NcPoint\_pop14  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NcLower\_pop14 NcUpper\_pop14 SourceNc\_pop14 Comments\_pop14 Name\_pop15  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## Origin\_pop15 IntroductionYear\_pop15 Ne\_pop15 NeLower\_pop15 NeUpper\_pop15  
## 1 NA NA NA  
## 2 NA NA NA  
## 3 NA NA NA  
## 4 NA NA NA  
## 5 NA NA NA  
## NeYear\_pop15 GeneticMarkers\_pop15 GeneticMarkersOther\_pop15 MethodNe\_pop15  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## SourceNe\_pop15 NcType\_pop15 NcYear\_pop15 NcMethod\_pop15 NcRange\_pop15  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcRangeDetails\_pop15 NcPoint\_pop15 NcLower\_pop15 NcUpper\_pop15 SourceNc\_pop15  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop15 Name\_pop16 Origin\_pop16 IntroductionYear\_pop16 Ne\_pop16  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NeLower\_pop16 NeUpper\_pop16 NeYear\_pop16 GeneticMarkers\_pop16  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## GeneticMarkersOther\_pop16 MethodNe\_pop16 SourceNe\_pop16 NcType\_pop16  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## NcYear\_pop16 NcMethod\_pop16 NcRange\_pop16 NcRangeDetails\_pop16 NcPoint\_pop16  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NcLower\_pop16 NcUpper\_pop16 SourceNc\_pop16 Comments\_pop16 Name\_pop17  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## Origin\_pop17 IntroductionYear\_pop17 Ne\_pop17 NeLower\_pop17 NeUpper\_pop17  
## 1 NA NA NA  
## 2 NA NA NA  
## 3 NA NA NA  
## 4 NA NA NA  
## 5 NA NA NA  
## NeYear\_pop17 GeneticMarkers\_pop17 GeneticMarkersOther\_pop17 MethodNe\_pop17  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## SourceNe\_pop17 NcType\_pop17 NcYear\_pop17 NcMethod\_pop17 NcRange\_pop17  
## 1   
## 2   
## 3   
## 4   
## 5   
## NcRangeDetails\_pop17 NcPoint\_pop17 NcLower\_pop17 NcUpper\_pop17 SourceNc\_pop17  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop17 Name\_pop18 Origin\_pop18 IntroductionYear\_pop18 Ne\_pop18  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NeLower\_pop18 NeUpper\_pop18 NeYear\_pop18 GeneticMarkers\_pop18  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## GeneticMarkersOther\_pop18 MethodNe\_pop18 SourceNe\_pop18 NcType\_pop18  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcYear\_pop18 NcMethod\_pop18 NcRange\_pop18 NcRangeDetails\_pop18 NcPoint\_pop18  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NcLower\_pop18 NcUpper\_pop18 SourceNc\_pop18 Comments\_pop18 Name\_pop19  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## Origin\_pop19 IntroductionYear\_pop19 Ne\_pop19 NeLower\_pop19 NeUpper\_pop19  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## NeYear\_pop19 GeneticMarkers\_pop19 GeneticMarkersOther\_pop19 MethodNe\_pop19  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## SourceNe\_pop19 NcType\_pop19 NcYear\_pop19 NcMethod\_pop19 NcRange\_pop19  
## 1   
## 2   
## 3   
## 4   
## 5   
## NcRangeDetails\_pop19 NcPoint\_pop19 NcLower\_pop19 NcUpper\_pop19 SourceNc\_pop19  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop19 Name\_pop20 Origin\_pop20 IntroductionYear\_pop20 Ne\_pop20  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NeLower\_pop20 NeUpper\_pop20 NeYear\_pop20 GeneticMarkers\_pop20  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## GeneticMarkersOther\_pop20 MethodNe\_pop20 SourceNe\_pop20 NcType\_pop20  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcYear\_pop20 NcMethod\_pop20 NcRange\_pop20 NcRangeDetails\_pop20 NcPoint\_pop20  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NcLower\_pop20 NcUpper\_pop20 SourceNc\_pop20 Comments\_pop20 Name\_pop21  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## Origin\_pop21 IntroductionYear\_pop21 Ne\_pop21 NeLower\_pop21 NeUpper\_pop21  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## NeYear\_pop21 GeneticMarkers\_pop21 GeneticMarkersOther\_pop21 MethodNe\_pop21  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## SourceNe\_pop21 NcType\_pop21 NcYear\_pop21 NcMethod\_pop21 NcRange\_pop21  
## 1   
## 2   
## 3   
## 4   
## 5   
## NcRangeDetails\_pop21 NcPoint\_pop21 NcLower\_pop21 NcUpper\_pop21 SourceNc\_pop21  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop21 Name\_pop22 Origin\_pop22 IntroductionYear\_pop22 Ne\_pop22  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NeLower\_pop22 NeUpper\_pop22 NeYear\_pop22 GeneticMarkers\_pop22  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## GeneticMarkersOther\_pop22 MethodNe\_pop22 SourceNe\_pop22 NcType\_pop22  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## NcYear\_pop22 NcMethod\_pop22 NcRange\_pop22 NcRangeDetails\_pop22 NcPoint\_pop22  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NcLower\_pop22 NcUpper\_pop22 SourceNc\_pop22 Comments\_pop22 Name\_pop23  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## Origin\_pop23 IntroductionYear\_pop23 Ne\_pop23 NeLower\_pop23 NeUpper\_pop23  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## NeYear\_pop23 GeneticMarkers\_pop23 GeneticMarkersOther\_pop23 MethodNe\_pop23  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## SourceNe\_pop23 NcType\_pop23 NcYear\_pop23 NcMethod\_pop23 NcRange\_pop23  
## 1 NA   
## 2 NA   
## 3 NA   
## 4 NA   
## 5 NA   
## NcRangeDetails\_pop23 NcPoint\_pop23 NcLower\_pop23 NcUpper\_pop23 SourceNc\_pop23  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## Comments\_pop23 Name\_pop24 Origin\_pop24 IntroductionYear\_pop24 Ne\_pop24  
## 1 NA NA  
## 2 NA NA  
## 3 NA NA  
## 4 NA NA  
## 5 NA NA  
## NeLower\_pop24 NeUpper\_pop24 NeYear\_pop24 GeneticMarkers\_pop24  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## GeneticMarkersOther\_pop24 MethodNe\_pop24 SourceNe\_pop24 NcType\_pop24  
## 1 NA NA NA   
## 2 NA NA NA   
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## NcYear\_pop24 NcMethod\_pop24 NcRange\_pop24 NcRangeDetails\_pop24 NcPoint\_pop24  
## 1 NA  
## 2 NA  
## 3 NA  
## 4 NA  
## 5 NA  
## NcLower\_pop24 NcUpper\_pop24 SourceNc\_pop24 Comments\_pop24 Name\_pop25  
## 1 NA NA   
## 2 NA NA   
## 3 NA NA   
## 4 NA NA   
## 5 NA NA   
## Origin\_pop25 IntroductionYear\_pop25 Ne\_pop25 NeLower\_pop25 NeUpper\_pop25  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## NeYear\_pop25 GeneticMarkers\_pop25 GeneticMarkersOther\_pop25 MethodNe\_pop25  
## 1 NA NA NA NA  
## 2 NA NA NA NA  
## 3 NA NA NA NA  
## 4 NA NA NA NA  
## 5 NA NA NA NA  
## SourceNe\_pop25 NcType\_pop25 NcYear\_pop25 NcMethod\_pop25 NcRange\_pop25  
## 1 NA NA NA  
## 2 NA NA NA  
## 3 NA NA NA  
## 4 NA NA NA  
## 5 NA NA NA  
## NcRangeDetails\_pop25 NcPoint\_pop25 NcLower\_pop25 NcUpper\_pop25 SourceNc\_pop25  
## 1 NA NA NA NA   
## 2 NA NA NA NA   
## 3 NA NA NA NA   
## 4 NA NA NA NA   
## 5 NA NA NA NA   
## Comments\_pop25 section6\_note nc\_type\_sp nc\_range\_sp nc\_range\_details\_sp  
## 1 NA   
## 2 NA Nc\_range less\_5000\_bymuch 3-21  
## 3 NA   
## 4 NA   
## 5 NA   
## nc\_point\_sp nc\_lower\_sp nc\_upper\_sp nc\_year\_sp Nc\_method\_sp  
## 1 NA NA NA   
## 2 NA NA NA 2014 Nc\_method\_count  
## 3 NA NA NA   
## 4 NA NA NA   
## 5 NA NA NA   
## source\_nc section6\_comments  
## 1   
## 2 https://doi.org/10.22201/fc.25942158e.2022.1.394   
## 3   
## 4   
## 5   
## section7\_note gen\_studies temp\_gen\_monitoring  
## 1 NA phylo\_pop yes  
## 2 NA no unknown  
## 3 NA phylo\_pop no  
## 4 NA phylo no  
## 5 NA phylo no  
## gen\_monitoring\_years  
## 1 Broader study: Historical samples from pre-1949 and modern from 1990-2012; Hluhluwe-iMfolozi 2001 and 2009-2014  
## 2   
## 3   
## 4   
## 5   
## source\_genetic\_studies  
## 1 Bertola LD, Jongbloed H, van der Gaag KJ, de Knijff P, Yamaguchi N, Hooghiemstra H, Bauer H, Henschel P, White PA, Driscoll CA, Tende T, Ottosson U, Saidu Y, Vrieling K, de Iongh HH. 2016. Phylogeographic patterns in Africa and high resolution delineation of genetic clades in the lion (Panthera leo). Scientific Reports 6: 30807; \nCurry CJ, Davis BW, Bertola LD, White PA, Murphy WJ, Derr JN. 2021. Spatiotemporal genetic diversity of lions reveals the influence of habitat fragmentation across Africa. Molecular Biology and Evolution 38(1), 48-57; \nDubach J, Patterson BD, Briggs MB, Venzke K, Flamand J, Stander P, Scheepers L, Kays RW. 2005. Molecular genetic variation across the southern and eastern geographic ranges of the African lion, Panthera leo. Conservation Genetics 6:15–24; \nDubach JM, Briggs MB, White PA, Ament BA, Patterson BD. 2013. Genetic perspectives on “Lion Conservation Units” in Eastern and Southern Africa. Conservation Genetics 14:741–755; \nMiller SM, Harper CK, Bloomer P, Hofmeyr J, Funston PJ. 2015. Fenced and Fragmented: Conservation value of managed metapopulations. PLoS One 10(12): e0144605; \nMiller SM, Druce DJ, Dalton DL, Harper CK, Kotze A, Packer C, Slotow R, Bloomer P. 2020. Genetic rescue of an isolated African lion population. Conservation Genetics 21: 41-53.  
## 2   
## 3   
## 4   
## 5   
## section8\_note realm  
## 1 NA terrestrial  
## 2 NA terrestrial  
## 3 NA freshwater terrestrial  
## 4 NA terrestrial  
## 5 NA freshwater  
## IUCN\_habitat  
## 1 savanna shrubland  
## 2 artificial\_terrestial forest\_woodland marine\_coastal\_supratidal  
## 3 artificial\_terrestial  
## 4 unknown  
## 5 wetland  
## other\_habitat  
## 1   
## 2 "Am (cálido húmedo) y Aw (Semicálido, templado húmedo), con una temperatura de 18º a 22ºC anuales (Rzedowski, 1994; CONABIO, 2001). Habita en un intervalo altitudinal que va de los 650 a 1350 m."-Naturalista.org  
## 3 creek, wetlands  
## 4   
## 5   
## national\_endemic transboundary\_type other\_explain country\_proportion  
## 1 no 1\_or\_more\_countries less\_than\_10  
## 2 yes   
## 3 yes   
## 4 yes   
## 5 yes   
## species\_range rarity occurrence\_extent occurrence\_area  
## 1 wide\_ranging not\_rare 41,745  
## 2 restricted rare\_recent 2183 km 2   
## 3 wide\_ranging not\_rare > 20 000 km2 0  
## 4 restricted rare\_natural   
## 5 restricted rare\_natural   
## pop\_fragmentation\_level  
## 1 moderate  
## 2 moderate  
## 3 severe  
## 4 none  
## 5 moderate  
## species\_range\_comments  
## 1   
## 2   
## 3 Its distribution is severely fragmented and there are continuous declines in the number of mature individuals, the extent and quality of its habitat around the Valley of Mexico.  
## 4   
## 5   
## global\_IUCN regional\_redlist  
## 1 vu vu  
## 2 en en  
## 3 not\_assessed en  
## 4 not\_assessed not\_assessed  
## 5 en not\_assessed  
## other\_assessment\_status  
## 1 Regional red list status (2016) Least Concern  
## 2 "...está protegida por la ley mexicana bajo la categoría de "Protección Especial" (Pr)."-IUCN  
## 3 there is no presence of other settlements of this species, in addition to the fact that it has been most recently evaluated for the IUCN Red List of Threatened Species in 2019  
## 4   
## 5 Endangered  
## other\_assessment\_name  
## 1   
## 2   
## 3 0  
## 4   
## 5 ECOS  
## source\_status\_distribution  
## 1 Miller S, Riggio J, Funston P, Power RJ, Williams V, Child MF. 2016. A conservation assessment of Panthera leo. In Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT, editors. The Red List of Mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa.  
## 2 https://www.iucnredlist.org/species/59289/53979925  
## 3 It is located in the Sierra de las Cruces, which is located to the west and south of the Valley of Mexico, in the states of Morelos, Mexico and Mexico City. Its range extends from Villa de Carbón south to the Lagunas de Zempoala. It occurs at elevations between 2,450 and 3,487 m asl (Woolrich-Pina et al. 2017). Its extent of occurrence (EOO) is 1564 km 2 , and it is estimated that the distribution represents a single location defined by threat;It lives to the west and south of the Valley of Mexico, in an altitude range of 2,700 to 3,200 m above sea level;It is found in the high mountains of central Mexico, southeast of the Federal District and northeast of Morelos.  
## 4 https://iefectividad.conanp.gob.mx/i-efectividad/CyEN/PN%20Izta-Popo/1%20CONTEXTO%20Y%20PLANEACION/9%20RECURSOS%20NATURALES/PropuestaHumedalCastillejas.pdf; https://enciclovida.mx/especies/153674-castilleja-tolucensis  
## 5 U.S. Fish and Wildlife Service. (2019) Species Status Assessment Report for the Slenderclaw Crayfish (Cambarus cracens).  
## section8\_note2 fecundity semelparous\_offpring reproductive\_strategy  
## 1 NA 3-5 NA sexual viviparous  
## 2 NA uknown NA oviparous  
## 3 NA 1000s NA sexual  
## 4 NA uknown NA sexual cross\_pollination  
## 5 NA uknown NA sexual  
## reproductive\_strategy\_other adult\_age\_data other\_reproductive\_strategy  
## 1 max\_lifespan maturity   
## 2 unknown   
## 3 unknown   
## 4 max\_lifespan maturity   
## 5 max\_lifespan   
## longevity\_max longevity\_median longevity\_maturity longevity\_age  
## 1 11\_25\_years 1\_5\_years   
## 2   
## 3   
## 4 unknown less\_than\_1\_year   
## 5 1\_5\_years   
## life\_history\_based\_on life\_history\_sp\_basedon  
## 1 species\_assessed   
## 2 species\_assessed   
## 3 species\_assessed   
## 4 species\_assessed   
## 5 close\_relative Cambaridae family  
## sources\_life\_history  
## 1 Skinner JD, Chimimba CT. 2005. The Mammals of the Southern African Subregion. 3rd edition. Cambridge University Press.  
## 2 https://www.naturalista.mx/taxa/66972-Pseudoeurycea-lineola  
## 3 Ajolote arroyero de la Sierra de las Cruces (Ambystoma altamirani). (s. f.). https://enciclovida.mx/especies/35117-ambystoma-altamirani\nSEMARNAT, 2018. Programa de Acción para la Conservación de\nlas Especies Ambystoma spp, SEMARNAT/CONANP, México\n(Año de edición 2018).\nResource not found. (s. f.). http://ri.uaemex.mx/bitstream/handle/20.500.11799/110035/Ecolog%C3%ADa+y+conocimiento+tradicional+de+Ambystoma+alta  
## 4 https://rdcu.be/c79UM; https://www.gbif.org/es/species/3734635; https://enciclovida.mx/especies/153674-castilleja-tolucensis; https://iefectividad.conanp.gob.mx/i-efectividad/CyEN/PN%20Izta-Popo/1%20CONTEXTO%20Y%20PLANEACION/9%20RECURSOS%20NATURALES/PropuestaHumedalCastillejas.pdf;  
## 5 U.S. Fish and Wildlife Service. (2019) Species Status Assessment Report for the Slenderclaw Crayfish (Cambarus cracens).  
## X\_id X\_uuid X\_submission\_time  
## 1 204546632 01aa2ec2-f7e0-4966-900b-1c4cf9e739cc 2022-11-29T12:16:25  
## 2 223206444 9b1b59e0-cffa-42b9-98b0-f0adbb34a34c 2023-03-06T07:38:11  
## 3 223372204 0d042a15-9c72-415a-b088-55c1a8a9f18b 2023-03-06T22:11:20  
## 4 226824792 f41dcf91-0adc-4c00-a6d7-986c44bea2c5 2023-03-21T23:33:46  
## 5 233980393 37576a4b-38da-4d28-941e-f6ec75c86741 2023-04-25T20:39:26  
## X\_validation\_status X\_notes X\_status X\_submitted\_by  
## 1 validation\_status\_not\_approved NA submitted\_via\_web   
## 2 validation\_status\_not\_approved NA submitted\_via\_web   
## 3 validation\_status\_not\_approved NA submitted\_via\_web   
## 4 validation\_status\_not\_approved NA submitted\_via\_web   
## 5 validation\_status\_not\_approved NA submitted\_via\_web   
## X\_\_version\_\_ X\_tags X\_index taxon  
## 1 vKMjGqNatUUs8xsWPMRVnq NA 44 Panthera leo  
## 2 vGbQECC4zBukxRQQxhGxLR NA 135 Pseudoeurycea Lineola  
## 3 vGbQECC4zBukxRQQxhGxLR NA 137 Ambystoma Altamirani 0  
## 4 vGbQECC4zBukxRQQxhGxLR NA 158 Castilleja Tolucensis  
## 5 vPo3KYKNtfzpQpF48EHdqe NA 456 Cambarus cracens

# omit those records from data:  
kobo\_clean<- kobo\_output %>%  
 filter(X\_validation\_status!="validation\_status\_not\_approved")

Filter out any sort of tests

# select likely columns to say "test"  
cols= c("name\_assessor", "email\_assessor", "genus", "species", "subspecies\_variety",  
 "scientific\_authority", "common\_name", "GBIF\_taxonID", "NCBI\_taxonID", "time\_populations")  
  
# check for "test" on any of them  
kobo\_clean %>%   
 filter(if\_any(all\_of(cols), ~ grepl("test", .))) %>%   
 select(country\_assessment, name\_assessor, genus, species)

## country\_assessment name\_assessor genus species  
## 1 united\_kingdom test test test  
## 2 united\_kingdom test test test  
## 3 united\_kingdom test test test  
## 4 united\_kingdom test test test

# filter them out of dataset  
kobo\_clean<- kobo\_clean %>%   
 filter(if\_any(all\_of(cols), ~ !grepl("test", .)))

## Number of populations

### Change -999 to NA

In the form, -999 was used to mark taxa with unknown number of extant populations. This was used because answering the question was mandatory, so leaving it blank wasn’t possible. We have to change -999 to NA.

For n extant populations:

kobo\_clean<- kobo\_clean %>%  
 mutate(n\_extant\_populations= na\_if(n\_extant\_populations, -999))

For n extinct populations:

kobo\_clean<- kobo\_clean %>%  
 mutate(n\_extint\_populations= na\_if(n\_extint\_populations, -999))

### Correct “negative populations”

Once -999 was replaced by NA there should be no negative number of populations (if they are, they are typos that need to be corrected).

Check for extant populations:

kobo\_clean %>%  
 filter(n\_extant\_populations<0) %>%  
 select(country\_assessment, taxon, name\_assessor, n\_extant\_populations, n\_extint\_populations)

## [1] country\_assessment taxon name\_assessor   
## [4] n\_extant\_populations n\_extint\_populations  
## <0 rows> (or 0-length row.names)

Check for extinct populations:

kobo\_clean %>%  
 filter(n\_extint\_populations<0) %>%  
 select(country\_assessment, taxon, name\_assessor, n\_extant\_populations, n\_extint\_populations)

## [1] country\_assessment taxon name\_assessor   
## [4] n\_extant\_populations n\_extint\_populations  
## <0 rows> (or 0-length row.names)

### Correct 999 populations

Show which species (if any) have 999 EXTINCT populations. **Should this be -999?**

Check for extinct populations:

kobo\_clean %>%  
 filter(n\_extint\_populations==999) %>%  
 select(country\_assessment, taxon, name\_assessor, n\_extant\_populations, n\_extint\_populations, end)

## [1] country\_assessment taxon name\_assessor   
## [4] n\_extant\_populations n\_extint\_populations end   
## <0 rows> (or 0-length row.names)

Check for extant populations:

kobo\_clean %>%  
 filter(n\_extant\_populations==999) %>%  
 select(country\_assessment, taxon, name\_assessor, n\_extant\_populations, n\_extint\_populations, end)

## [1] country\_assessment taxon name\_assessor   
## [4] n\_extant\_populations n\_extint\_populations end   
## <0 rows> (or 0-length row.names)

### Double check 0 (zero) extant populations

Show which taxa (if any) have 0 (zero) extant populations. **Is this correct? needs to be manually checked**

kobo\_clean %>%  
 filter(n\_extant\_populations==0) %>%  
 select(country\_assessment, taxon, name\_assessor, n\_extant\_populations, n\_extint\_populations, end)

## country\_assessment taxon name\_assessor  
## 1 france Hieracium sandozianum Myriam Heuertz  
## 2 australia Lasionectes exleyi Tanya Latty  
## 3 mexico Melanerpes aurifrons Alexander Llanes Quevedo  
## 4 united\_states Planorbella magnifica Meg Mahoney  
## 5 sweden Charadrius alexandrinus Viktoria Köppä  
## 6 united\_states Heterelmis stephani Taylor Stack  
## n\_extant\_populations n\_extint\_populations end  
## 1 0 1 2023-01-30T10:40:33.914+01:00  
## 2 0 NA 2023-04-06T17:34:39.716+10:00  
## 3 0 NA 2023-04-16T23:44:04.292-06:00  
## 4 0 3 2023-06-11T19:25:49.741-05:00  
## 5 0 3 2023-04-27T11:25:36.761+02:00  
## 6 0 2 2023-06-21T13:04:06.562-06:00

According to the assessors feedback the data of the following taxa is correct. For example, there are 0 extant populations because it is regionally extinct:

ok\_0\_extant<-c("Hieracium sandozianum",  
 "Charadrius alexandrinus",  
 "Planorbella magnifica")

Filtering out those taxa that are correct we should have an empity object, is not, the remaining taxa still need to be verified:

kobo\_clean %>%  
 filter(n\_extant\_populations==0) %>%  
 filter(taxon %!in% ok\_0\_extant) %>%  
 select(country\_assessment, taxon, name\_assessor, n\_extant\_populations, n\_extint\_populations, end)

## country\_assessment taxon name\_assessor  
## 1 australia Lasionectes exleyi Tanya Latty  
## 2 mexico Melanerpes aurifrons Alexander Llanes Quevedo  
## 3 united\_states Heterelmis stephani Taylor Stack  
## n\_extant\_populations n\_extint\_populations end  
## 1 0 NA 2023-04-06T17:34:39.716+10:00  
## 2 0 NA 2023-04-16T23:44:04.292-06:00  
## 3 0 2 2023-06-21T13:04:06.562-06:00

### and object with all “n pops” issues together

If there are any remaining taxa with weird number of populations put them together in a single object.

If everything is correct now, the object should be empity.

check\_n\_pops <- kobo\_clean %>%   
 # variables of interest  
 select(country\_assessment, name\_assessor, taxon, n\_extant\_populations, n\_extint\_populations, end) %>%  
  
 # same filters that discussed above  
 filter(n\_extant\_populations<0 |  
 n\_extant\_populations==0 |   
 n\_extant\_populations==999 |   
 n\_extint\_populations==999) %>%  
 filter(taxon %!in% ok\_0\_extant) %>%  
  
# add a column stating what needs to be checked:  
  
 mutate(need\_to\_check="check number of extant or extint populations. Are 0 correct? should 999 be -999? are extant/extint confused?")  
check\_n\_pops

## country\_assessment name\_assessor taxon  
## 1 australia Tanya Latty Lasionectes exleyi  
## 2 mexico Alexander Llanes Quevedo Melanerpes aurifrons  
## 3 united\_states Taylor Stack Heterelmis stephani  
## n\_extant\_populations n\_extint\_populations end  
## 1 0 NA 2023-04-06T17:34:39.716+10:00  
## 2 0 NA 2023-04-16T23:44:04.292-06:00  
## 3 0 2 2023-06-21T13:04:06.562-06:00  
## need\_to\_check  
## 1 check number of extant or extint populations. Are 0 correct? should 999 be -999? are extant/extint confused?  
## 2 check number of extant or extint populations. Are 0 correct? should 999 be -999? are extant/extint confused?  
## 3 check number of extant or extint populations. Are 0 correct? should 999 be -999? are extant/extint confused?

## Correct GBIF ids

The quality check script flagged any records where the GBIF Id is =/= 7, because that is the most common length. Changes after the assessors checked their data:

Corrections requested:

# Gymnobelideus leadbeateri   
  
kobo\_clean[kobo\_clean$taxon=="Gymnobelideus leadbeateri", "GBIF\_taxonID"]

## [1] "220657"

kobo\_clean[kobo\_clean$taxon=="Gymnobelideus leadbeateri", "GBIF\_taxonID"] <-2440054  
kobo\_clean[kobo\_clean$taxon=="Gymnobelideus leadbeateri", "GBIF\_taxonID"]

## [1] "2440054"

# Zingel asper  
kobo\_clean[kobo\_clean$taxon=="Zingel asper", "GBIF\_taxonID"]

## [1] "172830100"

kobo\_clean[kobo\_clean$taxon=="Zingel asper", "GBIF\_taxonID"] <- 2382117  
kobo\_clean[kobo\_clean$taxon=="Zingel asper", "GBIF\_taxonID"]

## [1] "2382117"

# Miniopterus schreibersii  
kobo\_clean[kobo\_clean$taxon=="Miniopterus schreibersii", "GBIF\_taxonID"]

## [1] "176678740"

kobo\_clean[kobo\_clean$taxon=="Miniopterus schreibersii", "GBIF\_taxonID"] <- 9796816  
kobo\_clean[kobo\_clean$taxon=="Miniopterus schreibersii", "GBIF\_taxonID"]

## [1] "9796816"

# Toxolasma lividum  
kobo\_clean[kobo\_clean$taxon=="Toxolasma lividum", "GBIF\_taxonID"]

## [1] "-999"

kobo\_clean[kobo\_clean$taxon=="Toxolasma lividum", "GBIF\_taxonID"] <-157572593  
kobo\_clean[kobo\_clean$taxon=="Toxolasma lividum", "GBIF\_taxonID"]

## [1] "157572593"

Correct GBIF ids even if they look weird:

ok\_GBIF<-c("Caladenia woolcockiorum",  
 "Leucopatus anophthalmus",   
 "Phyllurus kabikabi",  
 "Phonotimpus talquian")  
  
## add the manually corrected too, in case they are != 7  
  
ok\_GBIF<-c(ok\_GBIF,  
 "Gymnobelideus leadbeateri",  
 "Zingel asper",  
 "Miniopterus schreibersii",  
 "Toxolasma lividum")

### and object with all GBIF issues together

If there are any remaining taxa with weird GBIF ids put them together in a single object.

check\_GBIF <- kobo\_clean %>%  
 filter(nchar(GBIF\_taxonID)>0, nchar(GBIF\_taxonID)!=7) %>%  
 filter(taxon %!in% ok\_GBIF) %>%  
 # show only relevant columns  
 select(country\_assessment, name\_assessor, taxon, GBIF\_taxonID, end) %>%  
# add a column stating what needs to be checked:  
  
 mutate(need\_to\_check="check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).")  
check\_GBIF

## country\_assessment name\_assessor taxon  
## 1 south\_africa Samantha Mynhardt Chersobius signatus  
## 2 australia Catherine Grueber Hypotaenidia sylvestris  
## 3 australia Rebecca Jordan Ambuchanania leuchbryoides  
## 4 united\_states Brenna Forester Anaxyrus williamsi  
## 5 mexico Ana Laura Wegier Briuolo Gossypium davidsonii  
## 6 australia Georgina Wood Ecklonia radiata brevipes  
## 7 belgium Luis Castillo Schoenoplectus triqueter  
## GBIF\_taxonID end  
## 1 11030052 2023-05-24T10:44:48.654+02:00  
## 2 10913878 2023-05-03T01:00:28.585+10:00  
## 3 5792195; 8052578 2023-05-04T12:11:58.286+10:00  
## 4 10897495 2023-05-09T11:25:37.262-06:00  
## 5 112671-2 2023-05-24T11:37:32.648-06:00  
## 6 177596818 2023-06-03T12:06:02.619+02:00  
## 7 316510 2023-06-04T18:36:06.923+02:00  
## need\_to\_check  
## 1 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 2 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 3 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 4 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 5 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 6 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 7 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).

## Taxon names

Genus, species and subspecies should be a single word, check if there are cases where it isn’t. Only exception would be “var.” or “subsp.” in the subspecies\_variety field:

kobo\_clean %>%   
 filter(grepl(" ", genus) |   
 grepl(" ", species) |   
 grepl(" ", subspecies\_variety)) %>%  
 filter(!grepl("var.", subspecies\_variety)) %>%  
 filter(!grepl("subsp.", subspecies\_variety)) %>%  
 # show only relevant columns  
 # show only relevant columns  
 select(country\_assessment, name\_assessor, taxon, genus, species, subspecies\_variety, end)

## country\_assessment name\_assessor  
## 1 sweden Viktoria Köppä  
## 2 united\_states Lily Durkee  
## 3 mexico Alexander Llanes Quevedo  
## 4 mexico Alexander Llanes Quevedo  
## 5 mexico Alexander Llanes Quevedo  
## 6 mexico Alexander Llanes Quevedo  
## 7 mexico Orlando Jael Espinosa Chávez  
## 8 mexico Orlando Jael Espinosa Chávez  
## 9 mexico Alejandra Barrios Flores  
## taxon  
## 1 Coregonus albula morphotype trybomi  
## 2 Glyptemys muhlenbergii Northern Population  
## 3 Saltator atriceps atriceps, suffuscus, flavicrissus, peeti, raptor, lacertosus  
## 4 Icterus Icterus gularis gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas  
## 5 Melanerpes santacruzi grateloupensis, dubius, santacruzi, hughlandi, leei, turneffensis, pauper, insulanus, canescens  
## 6 Campylorhynchus rufinucha humilis, rufinucha, nigricaudatus, xerophilus, nicaraguae, castaneus, capistratus  
## 7 Leptotila verreauxi capitalis, angelica, bangsi and fulviventris  
## 8 Cyclarhys gujanensis septentrionalis, flaviventris, yucatanensis, insularis, nicaraguae  
## 9 Chirostoma estor estor; copandaro  
## genus species  
## 1 Coregonus albula  
## 2 Glyptemys muhlenbergii  
## 3 Saltator atriceps  
## 4 Icterus Icterus gularis  
## 5 Melanerpes santacruzi  
## 6 Campylorhynchus rufinucha  
## 7 Leptotila verreauxi  
## 8 Cyclarhys gujanensis  
## 9 Chirostoma estor  
## subspecies\_variety  
## 1 morphotype trybomi  
## 2 Northern Population  
## 3 atriceps, suffuscus, flavicrissus, peeti, raptor, lacertosus  
## 4 gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas  
## 5 grateloupensis, dubius, santacruzi, hughlandi, leei, turneffensis, pauper, insulanus, canescens  
## 6 humilis, rufinucha, nigricaudatus, xerophilus, nicaraguae, castaneus, capistratus  
## 7 capitalis, angelica, bangsi and fulviventris  
## 8 septentrionalis, flaviventris, yucatanensis, insularis, nicaraguae  
## 9 estor; copandaro  
## end  
## 1 2022-11-24T21:42:37.962+01:00  
## 2 2023-04-04T21:05:33.245-06:00  
## 3 2023-04-14T14:33:09.470-06:00  
## 4 2023-04-14T14:39:14.254-06:00  
## 5 2023-04-16T21:43:46.926-06:00  
## 6 2023-04-16T23:27:37.559-06:00  
## 7 2023-05-03T21:33:55.158-06:00  
## 8 2023-05-04T01:19:58.535-06:00  
## 9 2023-05-22T12:42:33.312-06:00

Corrections requested:

# Gypaetus Gypaetus barbatus  
kobo\_clean[kobo\_clean$taxon=="Gypaetus Gypaetus barbatus", "species"]

## character(0)

kobo\_clean[kobo\_clean$taxon=="Gypaetus Gypaetus barbatus", "species"] <-"barbatus"  
kobo\_clean[kobo\_clean$taxon=="Gypaetus Gypaetus barbatus", "taxon"] <-"Gypaetus barbatus"  
kobo\_clean[kobo\_clean$taxon=="Gypaetus barbatus", c("species", "taxon")]

## species taxon  
## 50 barbatus Gypaetus barbatus  
## 87 barbatus Gypaetus barbatus

# Gyps Gyps coprotheres  
kobo\_clean[kobo\_clean$taxon=="Gyps Gyps coprotheres", "species"]

## character(0)

kobo\_clean[kobo\_clean$taxon=="Gyps Gyps coprotheres", "species"] <-"coprotheres"  
kobo\_clean[kobo\_clean$taxon=="Gypaetus Gypaetus barbatus", "taxon"] <-"Gypaetus barbatus"  
kobo\_clean[kobo\_clean$taxon=="Gypaetus barbatus", c("species", "taxon")]

## species taxon  
## 50 barbatus Gypaetus barbatus  
## 87 barbatus Gypaetus barbatus

# Hippocampus Hippocampus capensis  
kobo\_clean[kobo\_clean$taxon=="Hippocampus Hippocampus capensis", "species"]

## character(0)

kobo\_clean[kobo\_clean$taxon=="Hippocampus Hippocampus capensis", "species"] <-"capensis"  
kobo\_clean[kobo\_clean$taxon=="Hippocampus Hippocampus capensis", "taxon"] <-"Hippocampus capensis"  
kobo\_clean[kobo\_clean$taxon=="Hippocampus capensis", c("species", "taxon")]

## species taxon  
## 52 capensis Hippocampus capensis

# Poicephalus Poicephalus robustus  
kobo\_clean[kobo\_clean$taxon=="Poicephalus Poicephalus robustus", "species"]

## character(0)

kobo\_clean[kobo\_clean$taxon=="Poicephalus Poicephalus robustus", "species"] <-"robustus"  
kobo\_clean[kobo\_clean$taxon=="Poicephalus Poicephalus robustus", "taxon"] <-"Poicephalus robustus"  
kobo\_clean[kobo\_clean$taxon=="Poicephalus robustus", c("species", "taxon")]

## species taxon  
## 54 robustus Poicephalus robustus

# Campylorhynchus rufinucha humilis, rufinucha, nigricaudatus, xerophilus, nicaraguae, castaneus, capistratus  
kobo\_clean[kobo\_clean$taxon=="Campylorhynchus rufinucha humilis, rufinucha, nigricaudatus, xerophilus, nicaraguae, castaneus, capistratus", "subspecies\_variety"]

## [1] "humilis, rufinucha, nigricaudatus, xerophilus, nicaraguae, castaneus, capistratus"

kobo\_clean[kobo\_clean$taxon=="Campylorhynchus rufinucha humilis, rufinucha, nigricaudatus, xerophilus, nicaraguae, castaneus, capistratus", "subspecies\_variety"] <-""  
kobo\_clean[kobo\_clean$taxon=="Campylorhynchus rufinucha humilis, rufinucha, nigricaudatus, xerophilus, nicaraguae, castaneus, capistratus", "taxon"] <- "Campylorhynchus rufinucha"  
kobo\_clean[kobo\_clean$taxon=="Campylorhynchus rufinucha", c("subspecies\_variety", "species", "taxon")]

## subspecies\_variety species taxon  
## 350 rufinucha Campylorhynchus rufinucha

# Icterus Icterus gularis gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas  
kobo\_clean[kobo\_clean$taxon=="Icterus Icterus gularis gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas", "subspecies\_variety"]

## [1] "gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas"

kobo\_clean[kobo\_clean$taxon=="Icterus Icterus gularis gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas", "subspecies\_variety"] <-""  
kobo\_clean[kobo\_clean$taxon=="Icterus Icterus gularis gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas", "species"]

## [1] "Icterus gularis"

kobo\_clean[kobo\_clean$taxon=="Icterus Icterus gularis gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas", "species"] <-"gularis"  
kobo\_clean[kobo\_clean$taxon=="Icterus Icterus gularis gularis, tamaulipensis, yucatanensis, flavescens, troglodytes, gigas", "taxon"] <- "Icterus gularis"  
kobo\_clean[kobo\_clean$taxon=="Icterus gularis", c("subspecies\_variety", "species", "taxon")]

## subspecies\_variety species taxon  
## 336 gularis Icterus gularis

# Melanerpes santacruzi grateloupensis, dubius, santacruzi, hughlandi, leei, turneffensis, pauper, insulanus, canescens  
kobo\_clean[kobo\_clean$taxon=="Melanerpes santacruzi grateloupensis, dubius, santacruzi, hughlandi, leei, turneffensis, pauper, insulanus, canescens", "subspecies\_variety"]

## [1] "grateloupensis, dubius, santacruzi, hughlandi, leei, turneffensis, pauper, insulanus, canescens"

kobo\_clean[kobo\_clean$taxon=="Melanerpes santacruzi grateloupensis, dubius, santacruzi, hughlandi, leei, turneffensis, pauper, insulanus, canescens", "subspecies\_variety"] <-""  
kobo\_clean[kobo\_clean$taxon=="Melanerpes santacruzi grateloupensis, dubius, santacruzi, hughlandi, leei, turneffensis, pauper, insulanus, canescens", "taxon"] <- "Melanerpes santacruzi"  
kobo\_clean[kobo\_clean$taxon=="Melanerpes santacruzi", c("subspecies\_variety", "species", "taxon")]

## subspecies\_variety species taxon  
## 344 santacruzi Melanerpes santacruzi

# Saltator atriceps atriceps, suffuscus, flavicrissus, peeti, raptor, lacertosus  
kobo\_clean[kobo\_clean$taxon=="Saltator atriceps atriceps, suffuscus, flavicrissus, peeti, raptor, lacertosus", "subspecies\_variety"]

## [1] "atriceps, suffuscus, flavicrissus, peeti, raptor, lacertosus"

kobo\_clean[kobo\_clean$taxon=="Saltator atriceps atriceps, suffuscus, flavicrissus, peeti, raptor, lacertosus", "subspecies\_variety"] <-""  
kobo\_clean[kobo\_clean$taxon=="Saltator atriceps atriceps, suffuscus, flavicrissus, peeti, raptor, lacertosus", "taxon"] <- "Saltator atriceps"  
kobo\_clean[kobo\_clean$taxon=="Saltator atriceps", c("subspecies\_variety", "species", "taxon")]

## subspecies\_variety species taxon  
## 335 atriceps Saltator atriceps

Correct:

ok\_taxon\_name<-c("Coregonus albula morphotype trybomi",  
 "Glyptemys muhlenbergii Northern Population")

### and object with all taxon names issues together

If there are any remaining taxa with weird taxon names put them together in a single object.

check\_taxon\_names <- kobo\_clean %>%   
 filter(grepl(" ", genus) |   
 grepl(" ", species) |   
 grepl(" ", subspecies\_variety)) %>%  
 filter(!grepl("var.", subspecies\_variety)) %>%  
 filter(!grepl("subsp.", subspecies\_variety)) %>%  
 filter(taxon %!in% ok\_taxon\_name) %>%  
 # show only relevant columns  
 select(country\_assessment, name\_assessor, taxon, genus, species, subspecies\_variety, end) %>%  
 mutate(need\_to\_check="check genus, species or subspecies\_variety, we are targeting to have single words in each field, except in the ifraspecific names, where 'var.' and 'subsp.' (only) would be accepted. Other details or taxonomic notes should be added in the comments.")  
check\_taxon\_names

## country\_assessment name\_assessor  
## 1 mexico Orlando Jael Espinosa Chávez  
## 2 mexico Orlando Jael Espinosa Chávez  
## 3 mexico Alejandra Barrios Flores  
## taxon  
## 1 Leptotila verreauxi capitalis, angelica, bangsi and fulviventris  
## 2 Cyclarhys gujanensis septentrionalis, flaviventris, yucatanensis, insularis, nicaraguae  
## 3 Chirostoma estor estor; copandaro  
## genus species  
## 1 Leptotila verreauxi  
## 2 Cyclarhys gujanensis  
## 3 Chirostoma estor  
## subspecies\_variety  
## 1 capitalis, angelica, bangsi and fulviventris  
## 2 septentrionalis, flaviventris, yucatanensis, insularis, nicaraguae  
## 3 estor; copandaro  
## end  
## 1 2023-05-03T21:33:55.158-06:00  
## 2 2023-05-04T01:19:58.535-06:00  
## 3 2023-05-22T12:42:33.312-06:00  
## need\_to\_check  
## 1 check genus, species or subspecies\_variety, we are targeting to have single words in each field, except in the ifraspecific names, where 'var.' and 'subsp.' (only) would be accepted. Other details or taxonomic notes should be added in the comments.  
## 2 check genus, species or subspecies\_variety, we are targeting to have single words in each field, except in the ifraspecific names, where 'var.' and 'subsp.' (only) would be accepted. Other details or taxonomic notes should be added in the comments.  
## 3 check genus, species or subspecies\_variety, we are targeting to have single words in each field, except in the ifraspecific names, where 'var.' and 'subsp.' (only) would be accepted. Other details or taxonomic notes should be added in the comments.

## Create a single file of pending issues for assessors review:

If there are still taxa that need to be checked for whatever reason, write a file enlisting them

to\_check<-full\_join(check\_n\_pops, check\_GBIF) %>% full\_join(check\_taxon\_names) %>%  
 # show columns in desired order:  
select(country\_assessment, name\_assessor, taxon, need\_to\_check, n\_extant\_populations,  
 n\_extint\_populations, GBIF\_taxonID, genus, species, subspecies\_variety, end)

## Joining, by = c("country\_assessment", "name\_assessor", "taxon", "end",  
## "need\_to\_check")  
## Joining, by = c("country\_assessment", "name\_assessor", "taxon", "end",  
## "need\_to\_check")

to\_check

## country\_assessment name\_assessor  
## 1 australia Tanya Latty  
## 2 mexico Alexander Llanes Quevedo  
## 3 united\_states Taylor Stack  
## 4 south\_africa Samantha Mynhardt  
## 5 australia Catherine Grueber  
## 6 australia Rebecca Jordan  
## 7 united\_states Brenna Forester  
## 8 mexico Ana Laura Wegier Briuolo  
## 9 australia Georgina Wood  
## 10 belgium Luis Castillo  
## 11 mexico Orlando Jael Espinosa Chávez  
## 12 mexico Orlando Jael Espinosa Chávez  
## 13 mexico Alejandra Barrios Flores  
## taxon  
## 1 Lasionectes exleyi  
## 2 Melanerpes aurifrons  
## 3 Heterelmis stephani  
## 4 Chersobius signatus  
## 5 Hypotaenidia sylvestris  
## 6 Ambuchanania leuchbryoides  
## 7 Anaxyrus williamsi  
## 8 Gossypium davidsonii  
## 9 Ecklonia radiata brevipes  
## 10 Schoenoplectus triqueter  
## 11 Leptotila verreauxi capitalis, angelica, bangsi and fulviventris  
## 12 Cyclarhys gujanensis septentrionalis, flaviventris, yucatanensis, insularis, nicaraguae  
## 13 Chirostoma estor estor; copandaro  
## need\_to\_check  
## 1 check number of extant or extint populations. Are 0 correct? should 999 be -999? are extant/extint confused?  
## 2 check number of extant or extint populations. Are 0 correct? should 999 be -999? are extant/extint confused?  
## 3 check number of extant or extint populations. Are 0 correct? should 999 be -999? are extant/extint confused?  
## 4 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 5 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 6 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 7 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 8 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 9 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 10 check the GBIF taxonID. Either it looks plain different, or has more or less than 7 digits (most ids are 7 digits long, and this isn't, it could be an exception, or a mistake).  
## 11 check genus, species or subspecies\_variety, we are targeting to have single words in each field, except in the ifraspecific names, where 'var.' and 'subsp.' (only) would be accepted. Other details or taxonomic notes should be added in the comments.  
## 12 check genus, species or subspecies\_variety, we are targeting to have single words in each field, except in the ifraspecific names, where 'var.' and 'subsp.' (only) would be accepted. Other details or taxonomic notes should be added in the comments.  
## 13 check genus, species or subspecies\_variety, we are targeting to have single words in each field, except in the ifraspecific names, where 'var.' and 'subsp.' (only) would be accepted. Other details or taxonomic notes should be added in the comments.  
## n\_extant\_populations n\_extint\_populations GBIF\_taxonID genus  
## 1 0 NA <NA> <NA>  
## 2 0 NA <NA> <NA>  
## 3 0 2 <NA> <NA>  
## 4 NA NA 11030052 <NA>  
## 5 NA NA 10913878 <NA>  
## 6 NA NA 5792195; 8052578 <NA>  
## 7 NA NA 10897495 <NA>  
## 8 NA NA 112671-2 <NA>  
## 9 NA NA 177596818 <NA>  
## 10 NA NA 316510 <NA>  
## 11 NA NA <NA> Leptotila  
## 12 NA NA <NA> Cyclarhys  
## 13 NA NA <NA> Chirostoma  
## species  
## 1 <NA>  
## 2 <NA>  
## 3 <NA>  
## 4 <NA>  
## 5 <NA>  
## 6 <NA>  
## 7 <NA>  
## 8 <NA>  
## 9 <NA>  
## 10 <NA>  
## 11 verreauxi  
## 12 gujanensis  
## 13 estor  
## subspecies\_variety  
## 1 <NA>  
## 2 <NA>  
## 3 <NA>  
## 4 <NA>  
## 5 <NA>  
## 6 <NA>  
## 7 <NA>  
## 8 <NA>  
## 9 <NA>  
## 10 <NA>  
## 11 capitalis, angelica, bangsi and fulviventris  
## 12 septentrionalis, flaviventris, yucatanensis, insularis, nicaraguae  
## 13 estor; copandaro  
## end  
## 1 2023-04-06T17:34:39.716+10:00  
## 2 2023-04-16T23:44:04.292-06:00  
## 3 2023-06-21T13:04:06.562-06:00  
## 4 2023-05-24T10:44:48.654+02:00  
## 5 2023-05-03T01:00:28.585+10:00  
## 6 2023-05-04T12:11:58.286+10:00  
## 7 2023-05-09T11:25:37.262-06:00  
## 8 2023-05-24T11:37:32.648-06:00  
## 9 2023-06-03T12:06:02.619+02:00  
## 10 2023-06-04T18:36:06.923+02:00  
## 11 2023-05-03T21:33:55.158-06:00  
## 12 2023-05-04T01:19:58.535-06:00  
## 13 2023-05-22T12:42:33.312-06:00

# save file with a list of taxa if there are remaining issues, or with a message saying that all is done if not):  
  
if(nrow(to\_check)>0) {  
  
write.csv(to\_check, "kobo\_output\_tocheck\_remaining\_issues.csv", row.names = FALSE, fileEncoding = "UTF-8")  
} else {  
happy\_message<-"congratulations! there are no more taxa with issues to be corrected. :)"  
 write.csv(happy\_message, "kobo\_output\_tocheck\_remaining\_issues.csv", row.names = FALSE, fileEncoding = "UTF-8")  
}

## Save the clean koboutput version:

Remove from the clean version any remaining taxa with issues

kobo\_clean<-kobo\_clean %>%  
 filter(taxon %!in% to\_check$taxon)

Export clean version

write.csv(kobo\_clean, "kobo\_output\_clean.csv", row.names = FALSE, fileEncoding = "UTF-8")

## Session Info for reproducibility purposes:

sessionInfo()

## R version 4.2.1 (2022-06-23)  
## Platform: x86\_64-apple-darwin17.0 (64-bit)  
## Running under: macOS Big Sur ... 10.16  
##   
## Matrix products: default  
## BLAS: /Library/Frameworks/R.framework/Versions/4.2/Resources/lib/libRblas.0.dylib  
## LAPACK: /Library/Frameworks/R.framework/Versions/4.2/Resources/lib/libRlapack.dylib  
##   
## locale:  
## [1] en\_US.UTF-8/en\_US.UTF-8/en\_US.UTF-8/C/en\_US.UTF-8/en\_US.UTF-8  
##   
## attached base packages:  
## [1] stats graphics grDevices utils datasets methods base   
##   
## other attached packages:  
## [1] ggplot2\_3.4.1 stringr\_1.4.0 utile.tools\_0.2.7 dplyr\_1.0.9   
## [5] tidyr\_1.2.0   
##   
## loaded via a namespace (and not attached):  
## [1] pillar\_1.7.0 compiler\_4.2.1 tools\_4.2.1 digest\_0.6.29   
## [5] evaluate\_0.15 lifecycle\_1.0.3 tibble\_3.1.7 gtable\_0.3.0   
## [9] pkgconfig\_2.0.3 rlang\_1.0.6 cli\_3.6.0 DBI\_1.1.3   
## [13] rstudioapi\_0.13 yaml\_2.3.5 xfun\_0.31 fastmap\_1.1.0   
## [17] withr\_2.5.0 knitr\_1.39 generics\_0.1.3 vctrs\_0.5.2   
## [21] grid\_4.2.1 tidyselect\_1.1.2 glue\_1.6.2 R6\_2.5.1   
## [25] fansi\_1.0.3 rmarkdown\_2.14 purrr\_0.3.4 magrittr\_2.0.3   
## [29] scales\_1.2.0 ellipsis\_0.3.2 htmltools\_0.5.5 assertthat\_0.2.1  
## [33] colorspace\_2.0-3 utf8\_1.2.2 stringi\_1.7.6 munsell\_0.5.0   
## [37] crayon\_1.5.1