Trees in Agriculture Policy Analysis

Lisa Meixner McCullough

2022-08-15

Table of Contents

[This is R Markdown 1](#_Toc114585521)

[Document Purpose 1](#_Toc114585522)

[Please note 2](#_Toc114585523)

[Initial Data Cleaning Protocol 2](#_Toc114585524)

[Convert data to wide format 2](#_Toc114585525)

[A view of the imported, pre-processed data 2](#_Toc114585526)

[Code to create a wide format document 3](#_Toc114585527)

[Code to create multiple long format documents 4](#_Toc114585528)

[Visualize Data 13](#_Toc114585529)

[Questions and Next Steps 20](#_Toc114585530)

[Graphing options 20](#_Toc114585531)

[I have three data questions: 20](#_Toc114585532)

# This is R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

# Document Purpose

In this workbook, I will  
1. Describe the pre-code data cleaning of the Trees in Agriculture policy recommendation data acquired by expert elicitation  
2. Reshape the data into wide form and long form and export these as csvs  
3. Use the long form data to visualize experts’ rankings of policy recommendations  
4. Outline questions and next steps

# Please note

* The csvs created by this code have been written to Lisa’s personal OneDrive.
  + Link to view: <https://conservation-my.sharepoint.com/:f:/g/personal/lmccullough_conservation_org/EuZlw83SDMhIj6wDobOJGVABM_oBhDYmZhalvVh2UnG7rA?e=d1L6HS>
  + The key documents (only ones you might want to use) in this folder are: -tia1\_policy\_wide.csv -tia`\_policy\_long.csv
    - you can open these in Excel by double clicking on them

# Initial Data Cleaning Protocol

Pre-upload to R, the data was cleaned in Excel using the following protocol:  
1. I renamed columns to be simple variable descriptions as opposed to full length question  
2. I removed the following columns because they were not important to data analysis: start time, completion time, choice to accept/waive honorarium, email, authorization of use  
3. I replaced all commas, in the entire document, with semi-colons. This is so commas within cells would not cause parsing errors when sheet was in csv format.  
4. For any ranking question, I recoded information from the ranked and manual entry into two columns: ranked Important and ranked Not important. I used the following protocol:  
- If an expert manually input a number that was not an option, e.g. “10”, I deleted these instances (rather than assuming they meant “1”, etc). - Some people manually input other reasons; I simplified these to “Other” and put “Other” in the list of important reasons immediately following the most closely relevant numbered suggestion. - If people input both in ranking entry and in manual entry, I used the manual entry and disregarded the ranking entry, even if they conflicted. - For any items “not important” I listed them in the order ranked, though the ranking is not technically relevant. - If manual entry was chosen, but not all items were listed, I assumed the unlisted items were “not important” and recorded them thusly. - If manual entry was used, and both important and unimportant items were listed, but the list was not complete, I categorized other items based on the ranked list, if the ranked list otherwise matched the manual list. - If an expert made all reasons important, then they therefore thought “None” were unimportant. (Thus language used). If the expert did not answer the question, then NA is used.

# Convert data to wide format

#read csv  
policy\_raw<-read\_delim("TIA\_Policy.csv",delim=",",col\_names=T)

## A view of the imported, pre-processed data

#knitr::kable(policy\_raw[1:5,],caption='Following edits made in Excel, this is the data that was imported to R')

## Code to create a wide format document

1. Use the tidyverse function separate() to separate out each ranked reason or policy into its own column.

#prepare to store output as policy\_wide  
policy\_wide<-  
 #start with raw data  
 policy\_raw %>%  
   
 #separate columns: Why Trees in Corn Boundaries?  
 separate(cornBImp,into=c("CBP1","CBP2","CBP3","CBP4","CBP5","CBP6","CBP7","CBP8","CBP9"),sep=";")%>%  
 separate(cornBNImp,into=c("CBN1","CBN2","CBN3","CBN4","CBN5","CBN6","CBN7","CBN8","CBN9"),sep=";")%>%  
   
 #separate columns: Why Trees in Corn Fields?  
 separate(cornFImp,into=c("CFP1","CFP2","CFP3","CFP4","CFP5","CFP6","CFP7","CFP8","CFP9"),sep=";")%>%  
 separate(cornFNImp,into=c("CFN1","CFN2","CFN3","CFN4","CFN5","CFN6","CFN7","CFN8","CFN9"),sep=";")%>%  
   
 #separate columns: Why Trees in Wheat Boundaries?  
 separate(wheatBImp,into=c("WBP1","WBP2","WBP3","WBP4","WBP5","WBP6","WBP7","WBP8","WBP9"),sep=";")%>%  
 separate(wheatBNImp,into=c("WBN1","WBN2","WBN3","WBN4","WBN5","WBN6","WBN7","WBN8","WBN9"),sep=";")%>%  
   
 #separate columns: Why Trees in Wheat Fields?  
 separate(wheatFImp,into=c("WFP1","WFP2","WFP3","WFP4","WFP5","WFP6","WFP7","WFP8","WFP9"),sep=";")%>%  
 separate(wheatFNImp,into=c("WFN1","WFN2","WFN3","WFN4","WFN5","WFN6","WFN7","WFN8","WFN9"),sep=";")%>%  
   
 #separate columns: Why Trees in Soy Boundaries?  
 separate(soyBImp,into=c("SBP1","SBP2","SBP3","SBP4","SBP5","SBP6","SBP7","SBP8","SBP9"),sep=";")%>%  
 separate(soyBNImp,into=c("SBN1","SBN2","SBN3","SBN4","SBN5","SBN6","SBN7","SBN8","SBN9"),sep=";")%>%  
   
 #separate columns: Why Trees in Soy Fields?  
 separate(soyFImp,into=c("SFP1","SFP2","SFP3","SFP4","SFP5","SFP6","SFP7","SFP8","SFP9"),sep=";")%>%  
 separate(soyFNImp,into=c("SFN1","SFN2","SFN3","SFN4","SFN5","SFN6","SFN7","SFN8","SFN9"),sep=";")%>%  
   
 #separate columns: What policy recommendations would help increase tree density?  
 separate(policiesImp,into=c("A1","A2","A3","A4","A5","A6"),sep=";")%>%  
 separate(policiesNImp,into=c("AN1","AN2","AN3","AN4","AN5","AN6"),sep=";")

1. Export reshaped, wide data to working directory as a csv named tia1\_policy\_wide.csv

#write.csv(policy\_wide,file="tia1\_policy\_wide.csv")

## Code to create multiple long format documents

Use the tidyverse function reshape to convert the wide format dataset to its long form.

1. Name which columns are linked to each data question using base function grep and regular expressions.

#CORN  
#corn boundary columns - important reasons  
cornB\_cols<-grep("CBP.",x=colnames(policy\_wide),value=T)  
#corn boundary columns - not important reasons  
cornNB\_cols<-grep("CBN.",x=colnames(policy\_wide),value=T)  
#corn field columns - important reasons  
cornF\_cols<-grep("CFP.",x=colnames(policy\_wide),value=T)  
#corn field columns - not important reasons  
cornNF\_cols<-grep("CFN.",x=colnames(policy\_wide),value=T)  
  
#WHEAT  
#wheat boundary columns - important reasons  
wheatB\_cols<-grep("WBP.",x=colnames(policy\_wide),value=T)  
#wheat boundary columns - not important reasons  
wheatNB\_cols<-grep("WBN.",x=colnames(policy\_wide),value=T)  
#wheat field columns - important reasons  
wheatF\_cols<-grep("WFP.",x=colnames(policy\_wide),value=T)  
#wheat field columns - not important reasons  
wheatNF\_cols<-grep("WFN.",x=colnames(policy\_wide),value=T)  
  
#SOY  
#soy boundary columns - important reasons  
soyB\_cols<-grep("SBP.",x=colnames(policy\_wide),value=T)  
#soy boundary columns - not important reasons  
soyNB\_cols<-grep("SBN.",x=colnames(policy\_wide),value=T)  
#soy field columns - important reasons  
soyF\_cols<-grep("SFP.",x=colnames(policy\_wide),value=T)  
#soy field columns - not important reasons  
soyNF\_cols<-grep("SFN.",x=colnames(policy\_wide),value=T)  
  
#POLICY  
#policy columns - important policies  
policy\_cols<-grep("A[0-9]",x=colnames(policy\_wide),value=T)  
#policy columns - not important policies  
policyN\_cols<-grep("AN[0-9]",x=colnames(policy\_wide),value=T)  
  
#Store all of these groups into two lists  
b\_cols<-data.frame(cornB\_cols,cornNB\_cols,wheatB\_cols,wheatNB\_cols,soyB\_cols,soyNB\_cols)  
f\_cols<-data.frame(cornF\_cols,cornNF\_cols,wheatF\_cols,wheatNF\_cols,soyF\_cols,soyNF\_cols)  
p\_cols<-data.frame(policy\_cols,policyN\_cols)

1. Describe Variables for Tidy / Long form data

*There are five meaningful variables to be included in the long form/tidy dataset that are hidden within the wide dataset.*

* crop an unordered factor with three levels: soy,wheat,corn
* boundaryReason an unordered factor with nine levels: acts as barrier or living fence, acts as windbreak, produces something of extractive value, improves microclimate, improves soil, provides habitat, aesthetics, farmers compensated, trees already present, which correspond to the following:
  + 1 The trees are a barrier or living fence or mark the property or field boundary
  + 2 The trees are windbreaks
  + 3 The trees are producing something of extractive value (ie. timber, fuelwood, fruit, animal forage, other)
  + 4 The trees improve the microclimate (provide shade, sunlight or temperature control, water retention, etc.)
  + 5 The trees improve the soil (retention, nutrients, etc.)
  + 6 The trees provide habitat (pollinators, beneficial predators, beneficial biodiversity)
  + 7 The trees are visually pleasing
  + 8 Farmers are compensated for the presence of trees as a result of a policy
  + 9 The trees were already present and were not removed
* fieldReason an unordered factor with nine levels: produces something of extractive value, acts as windbreak, improves microclimate, improves soil, provides habitat, aesthetics, farmers compensated, trees already present, other cultural/historical reason, which correspond to the following:
  + 1 The trees are producing something of extractive value (ie. timber, fuelwood, fruit, animal forage, other)
  + 2 The trees are windbreaks
  + 3 The trees improve the microclimate (provide shade, sunlight or temperature control, water retention, etc.)
  + 4 The trees improve the soil (retention, nutrients, etc.)
  + 5 The trees provide habitat (pollinators, beneficial predators, beneficial biodiversity)
  + 6 The trees are visually pleasing
  + 7 Farmers are compensated for the presence of trees as a result of a policy
  + 8 The trees were already present and were not removed
  + 9 Other historical or cultural practice not listed above
* policy an unordered factor with six levels: research,develop shade-grown variatals,develop/reform policy,provision inputs,expand ag extension services,expand carbon crediting/verification, which correspond to:
  + Research on tree-crop interactions
  + Development of shade-tolerant crop varieties
  + Policy development and/or reform
  + Provisioning of inputs
  + Agricultural extension services for technical assistance
  + Carbon crediting/certification methods for these systems
  + *Note:* policy is relevant to biome-continent pair and all cropland in this region. However, if doing higher level analysis (at biome-continent level), we should be wary of double counting. Unchecked, experts that responded for more than one crop would have their policy recommendations counted multiple times.
* priority an ordered factor with ten levels: First, Second, Third, Fourth, Fifth, Sixth, Seventh, Eighth, Ninth, Not Meaningful.
  + all reasons or policies that the experts called not important would receive the rank Not Meaningful
  + boundaryReason and fieldReason can use any of these levels; policyRecommendation uses only Not Meaningful and levels First through Sixth.

*There are seven other meaningful variables that do not change from wide to long/tidy data:*

* relevant for quantitative analysis:
  + continent, an unordered factor with five levels: Africa, Asia (includes Middle East), Europe, North/Central America (includes Caribbean) and South America
  + biome, an unordered factor with seven levels: Boreal Forests & Taiga, Mediterranean Forests; Woodland; & Scrub, Temperate Broadleaf & Mixed Forests, Tropical & Subtropical Dry Broadleaf Forests, Tropical & Subtropical Moist Broadleaf Forests,Temperate Grasslands; Savannas & Shrublands, Tropical & Subtropical Grasslands; Savannas; and Shrublands
* relevant for qualitative analysis: repTree,countryExpertise
  + repTree and countryExpertise would require cleaning prior to quantitative analysis, generally the entries are unique to the expert
* relevant for IDing instances of completing questionnaire: id
  + alas, one expert (Pierluigi) completed questionnaire three times, for three different tree species. so name combined with repTree can also ID a unique questionnaire response, but name alone cannot
  + institution provides more info about expert, but we want to avoid assumption that experts could not share the same institution

crop<-c("corn","wheat","soy")  
b\_or\_f<-c("boundaries","fields")  
  
reason<-c("acts as barrier or living fence", "acts as windbreak", "produces something of extractive value", "improves microclimate", "improves soil", "provides habitat", "aesthetics", "farmers compensated", "trees already present", "other cultural/historical reason")  
reasonImportance<-c("Top","Second","Third","Fourth","Fifth","Sixth","Seventh","Eighth","Last","Not Meaningful", "Not Meaningful","Not Meaningful","Not Meaningful","Not Meaningful","Not Meaningful","Not Meaningful","Not Meaningful","Not Meaningful")  
  
policy<-c("research","develop shade-grown variatals","develop/reform policy","provision inputs","expand ag extension services","expand carbon crediting/verification")  
policyPriority<-c("Top","Second","Third","Fourth","Fifth","Last","Not Meaningful","Not Meaningful", "Not Meaningful","Not Meaningful","Not Meaningful","Not Meaningful")  
  
general\_cols<-c("continent","biome","repTree","name","id","institution","countryExpertise")

1. Subset dataframes

#subset trees in boundaries data  
wide\_wb<-subset(policy\_wide,select=c(general\_cols,wheatB\_cols,wheatNB\_cols),value=T)  
wide\_cb<-subset(policy\_wide,select=c(general\_cols,cornB\_cols,cornNB\_cols),value=T)  
wide\_sb<-subset(policy\_wide,select=c(general\_cols,soyB\_cols,soyNB\_cols),value=T)  
  
#subset trees in fields data  
wide\_wf<-subset(policy\_wide,select=c(general\_cols,wheatF\_cols,wheatNF\_cols),value=T)  
wide\_cf<-subset(policy\_wide,select=c(general\_cols,cornF\_cols,cornNF\_cols),value=T)  
wide\_sf<-subset(policy\_wide,select=c(general\_cols,soyF\_cols,soyNF\_cols),value=T)  
  
#subset policy data  
wide\_p<-subset(policy\_wide,select=c(general\_cols,policy\_cols,policyN\_cols),value=T)

1. Create Long Formats of Subset dataframes

#CREATE LONG FORMATS OF SUBSET DATAFRAMES  
#boundary data frames  
#corn   
long\_cb<-  
 wide\_cb%>%  
 #reshape this chosen data frame from wide to long  
 reshape(  
 direction="long",  
 varying=c(cornB\_cols,cornNB\_cols),  
 v.name="reason",  
 timevar="reasonImportance",  
 idvar="id",  
 times=reasonImportance,  
 new.row.names=1:1000  
 )%>%  
 #create crop column and b\_or\_f column and code with appropriate values  
 #and recode reason column to use informative names (that don't appear ordered)  
 mutate(  
 crop="corn",  
 b\_or\_f="boundary",  
 reason=  
 recode(reason,  
 "1"="acts as barrier or living fence",   
 "2"="acts as windbreak",   
 "3"="produces something of extractive value",   
 "4"="improves microclimate",   
 "5"="improves soil",   
 "6"="provides habitat",   
 "7"="aesthetics",   
 "8"="farmers compensated",   
 "9"="trees already present")  
 )%>%  
 #in order to reduce object size, remove instances of reason = NA  
 dplyr::filter(!is.na(reason))  
#wheat  
long\_wb<-  
 wide\_wb%>%  
 #reshape this chosen data frame from wide to long  
 reshape(  
 direction="long",  
 varying=c(wheatB\_cols,wheatNB\_cols),  
 v.name="reason",  
 timevar="reasonImportance",  
 idvar="id",  
 times=reasonImportance,  
 new.row.names=1:1000  
 )%>%  
 mutate(  
 crop="wheat",  
 b\_or\_f="boundary",  
 reason=  
 recode(reason,  
 "1"="acts as barrier or living fence",   
 "2"="acts as windbreak",   
 "3"="produces something of extractive value",   
 "4"="improves microclimate",   
 "5"="improves soil",   
 "6"="provides habitat",   
 "7"="aesthetics",   
 "8"="farmers compensated",   
 "9"="trees already present")  
 )%>%  
 #in order to reduce object size, remove instances of reason = NA  
 dplyr::filter(!is.na(reason))  
#soy  
long\_sb<-  
 wide\_sb%>%  
 #reshape this chosen data frame from wide to long  
 reshape(  
 direction="long",  
 varying=c(soyB\_cols,soyNB\_cols),  
 v.name="reason",  
 timevar="reasonImportance",  
 idvar="id",  
 times=reasonImportance,  
 new.row.names=1:1000  
 )%>%  
 mutate(  
 crop="soy",  
 b\_or\_f="boundary",  
 reason=  
 recode(reason,  
 "1"="acts as barrier or living fence",   
 "2"="acts as windbreak",   
 "3"="produces something of extractive value",   
 "4"="improves microclimate",   
 "5"="improves soil",   
 "6"="provides habitat",   
 "7"="aesthetics",   
 "8"="farmers compensated",   
 "9"="trees already present")  
 )%>%  
 #in order to reduce object size, remove instances of reason = NA  
 dplyr::filter(!is.na(reason))  
  
#field data frames  
#corn  
long\_cf<-  
 wide\_cf%>%  
 #reshape this chosen data frame from wide to long  
 reshape(  
 direction="long",  
 varying=c(cornF\_cols,cornNF\_cols),  
 v.name="reason",  
 timevar="reasonImportance",  
 idvar="id",  
 times=reasonImportance,  
 new.row.names=1:1000  
 )%>%  
 mutate(  
 crop="corn",  
 b\_or\_f="field",  
 reason=  
 recode(reason,  
 "1"="acts as windbreak",   
 "2"="produces something of extractive value",   
 "3"="improves microclimate",   
 "4"="improves soil",   
 "5"="provides habitat",   
 "6"="aesthetics",   
 "7"="farmers compensated",   
 "8"="trees already present",  
 "9"="other cultural/historical reason")  
 )%>%  
 #in order to reduce object size, remove instances of reason = NA  
 dplyr::filter(!is.na(reason))  
#wheat  
long\_wf<-  
 wide\_wf%>%  
 #reshape this chosen data frame from wide to long  
 reshape(  
 direction="long",  
 varying=c(wheatF\_cols,wheatNF\_cols),  
 v.name="reason",  
 timevar="reasonImportance",  
 idvar="id",  
 times=reasonImportance,  
 new.row.names=1:1000  
 )%>%  
 mutate(  
 crop="wheat",  
 b\_or\_f="field",  
 reason=  
 recode(reason,  
 "1"="acts as windbreak",   
 "2"="produces something of extractive value",   
 "3"="improves microclimate",   
 "4"="improves soil",   
 "5"="provides habitat",   
 "6"="aesthetics",   
 "7"="farmers compensated",   
 "8"="trees already present",  
 "9"="other cultural/historical reason")  
 )%>%  
 #in order to reduce object size, remove instances of reason = NA  
 dplyr::filter(!is.na(reason))  
#soy  
long\_sf<-  
 wide\_sf%>%  
 #reshape this chosen data frame from wide to long  
 reshape(  
 direction="long",  
 varying=c(soyF\_cols,soyNF\_cols),  
 v.name="reason",  
 timevar="reasonImportance",  
 idvar="id",  
 times=reasonImportance,  
 new.row.names=1:1000  
 )%>%  
 mutate(  
 crop="soy",  
 b\_or\_f="field",  
 reason=  
 recode(reason,  
 "1"="acts as windbreak",   
 "2"="produces something of extractive value",   
 "3"="improves microclimate",   
 "4"="improves soil",   
 "5"="provides habitat",   
 "6"="aesthetics",   
 "7"="farmers compensated",   
 "8"="trees already present",  
 "9"="other cultural/historical reason")  
 )%>%  
 #in order to reduce object size, remove instances of reason = NA  
 dplyr::filter(!is.na(reason))  
#policy data frames  
long\_p<-  
 wide\_p%>%  
 #reshape this chosen data frame from wide to long  
 reshape(  
 direction="long",  
 varying=c(policy\_cols,policyN\_cols),  
 v.name="policy",  
 timevar="policyPriority",  
 idvar="id",  
 times=policyPriority,  
 new.row.names=1:1000  
 )%>%  
 mutate(  
 policy=recode(policy,  
 "1"="research",  
 "2"="develop shade-grown variatals",  
 "3"="develop/reform policy",  
 "4"="provision inputs",  
 "5"="expand ag extension services",  
 "6"="expand carbon crediting/verification")  
 )%>%  
 #in order to reduce object size, remove instances of reason = NA  
 dplyr::filter(!is.na(policy))

1. Write long dataframe to folder as csv and subsetted versions of this to subfolder tia1\_long\_data

#clear environment to free up space  
env\_all<-ls() #returns names of all objects  
env\_not\_needed<-env\_all[! env\_all %in% c("general\_cols","long\_p","long\_cb","long\_wb","long\_sb","long\_cf","long\_wf","long\_sf")] #only objects we don't need anymore  
rm(list=env\_not\_needed) #removes objects we don't need anymore  
  
policy\_long<-  
 long\_cb%>%  
 full\_join(long\_wb)%>%  
 full\_join(long\_sb)%>%  
 full\_join(long\_cf)%>%  
 full\_join(long\_wf)%>%  
 full\_join(long\_sf)%>%  
 full\_join(long\_p,by=general\_cols)%>%  
 mutate\_at(vars(reasonImportance),factor,levels=c("Top","Second","Third","Fourth","Fifth","Sixth","Seventh","Eighth","Last","Not Meaningful"))%>%  
 mutate\_at(vars(policyPriority),factor,levels=c("Not Meaningful","Last","Fifth","Fourth","Third","Second","Top"))  
write.csv(policy\_long, file="tia1\_policy\_long.csv")  
  
#write csvs of subsetted long data  
setwd("./tia1\_long\_data")  
write.csv(long\_p, file="tia1\_policyRecs\_long.csv")  
write.csv(long\_cb, file="tia1\_cornBound\_long.csv")  
write.csv(long\_wb, file="tia1\_wheatBound\_long.csv")  
write.csv(long\_sb, file="tia1\_soyBound\_long.csv")  
write.csv(long\_cf, file="tia1\_cornField\_long.csv")  
write.csv(long\_wf, file="tia1\_wheatField\_long.csv")  
write.csv(long\_sf, file="tia1\_soyField\_long.csv")

# Visualize Data

*For now this only shows choices for crop, with other defaults applied* - Defaults - It will plot “Reasons for Trees in Boundaries or in Fields” - It will include responses for all continents and all biomes

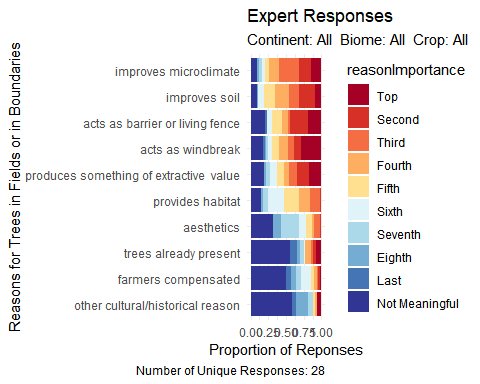
1. Create Histogram Function

crop\_histogram<-function(crop,data,what2plot,continent,biome){  
 #filter data to just crop requested   
 if(crop!="All"){  
 if(crop=="corn"){  
 data<-data%>%filter(crop=="corn")  
 }  
 if(crop=="wheat"){  
 data<-data%>%filter(crop=="wheat")  
 }  
 if(crop=="soy"){  
 data<-data%>%filter(crop=="soy")  
 }  
 }  
   
 #calc number of unique responses  
 dimDistinct<-transmute(data,name=name,repTree=repTree)%>%distinct()%>%dim()  
 numResp<-dimDistinct[1]  
   
 #reorder reason based on frequency  
 data<-data%>%  
 mutate(reasonValue=recode(reasonImportance,"Not Meaningful"=0,"Last"=1,"Eighth"=2,"Seventh"=3,"Sixth"=4,"Fifth"=5,"Fourth"=6,"Third"=7,"Second"=8,"Top"=9))  
print(data)  
print(unique(data$reason))  
  
order<-array(data=NA,dim=c(length(unique(data$reason)),4), dimnames=list(NULL,c("reason","sum.b","sum.f","rank")))  
n=1  
for(r in unique(data$reason)){  
 data.rb<-data%>%  
 filter(reason==r&b\_or\_f=="boundary")  
 data.rf<-data%>%  
 filter(reason==r&b\_or\_f=="field")  
 order[n,]<-c(r,  
 sum(data.rb$reasonValue),  
 sum(data.rf$reasonValue),  
 ifelse(min(c(sum(data.rb$reasonValue),sum(data.rf$reasonValue)))==0,  
 sum(sum(data.rb$reasonValue),sum(data.rf$reasonValue)),  
 mean(c(sum(data.rb$reasonValue),sum(data.rf$reasonValue)))  
 )  
 )  
 n=n+1  
}  
  
order<-order%>%  
 as.data.frame()%>%  
 mutate(rank=as.numeric(rank))%>%  
 arrange(rank)  
  
reasonOrder<-order$reason  
  
data<-data%>%  
 mutate\_at(vars(reason),factor,levels=reasonOrder)%>%  
 filter(reason!="Other"&reason!="None"&!is.na(reason))  
  
#display graph  
 data%>%  
 ggplot(mapping=aes(x=reason)) +  
 geom\_bar(position="fill",aes(fill=reasonImportance))+  
 coord\_flip()+  
 labs(y="Proportion of Reponses",x=what2plot,title="Expert Responses",subtitle=paste("Continent:",continent," Biome:",biome," Crop:",crop),caption=paste("Number of Unique Responses:",as.character(numResp)))+  
 # geom\_text(aes(y=1,label=b\_or\_f,hjust="right"),  
 # check\_overlap = T,  
 # nudge\_y = 0.2) +  
 #guides(color = guide\_legend(reverse=TRUE))+  
 theme\_minimal()+  
 scale\_fill\_brewer(type="div",palette="RdYlBu",direction=1,aesthetics=c("fill"))  
}

1. Display plot for All Crops

#initialize defaults  
what2plot<-"Reasons for Trees in Fields or in Boundaries"  
continent<-"All"  
biome<-"All"  
  
  
crop\_histogram(crop="All",data=policy\_long,what2plot=what2plot,continent=continent,biome=biome)

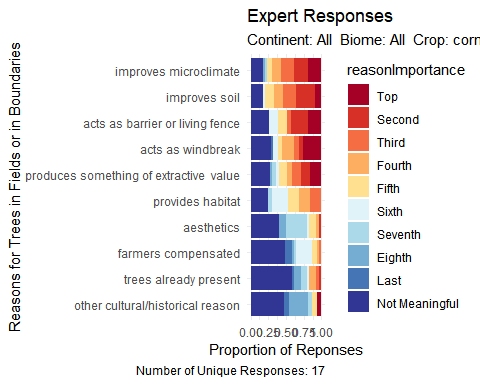
## # A tibble: 3,967 x 14  
## continent biome repTree name id institution countryExpertise  
## <chr> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 2 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 3 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 4 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 5 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 6 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 7 Africa Tropical & Subtro~ Alnus ~ Alai~ 5 Internatio~ Rwanda; Ethiopi~  
## 8 Africa Tropical & Subtro~ Alnus ~ Alai~ 5 Internatio~ Rwanda; Ethiopi~  
## 9 Africa Tropical & Subtro~ Alnus ~ Alai~ 5 Internatio~ Rwanda; Ethiopi~  
## 10 Africa Tropical & Subtro~ Alnus ~ Alai~ 5 Internatio~ Rwanda; Ethiopi~  
## # ... with 3,957 more rows, and 7 more variables: reasonImportance <fct>,  
## # reason <chr>, crop <chr>, b\_or\_f <chr>, policyPriority <fct>, policy <chr>,  
## # reasonValue <dbl>  
## [1] "acts as barrier or living fence"   
## [2] "produces something of extractive value"  
## [3] "Other"   
## [4] "improves soil"   
## [5] "improves microclimate"   
## [6] "acts as windbreak"   
## [7] "aesthetics"   
## [8] "trees already present"   
## [9] "provides habitat"   
## [10] "farmers compensated"   
## [11] "other cultural/historical reason"   
## [12] NA



1. Display plot for Corn

crop\_histogram(crop="corn",data=policy\_long,what2plot=what2plot,continent=continent,biome=biome)

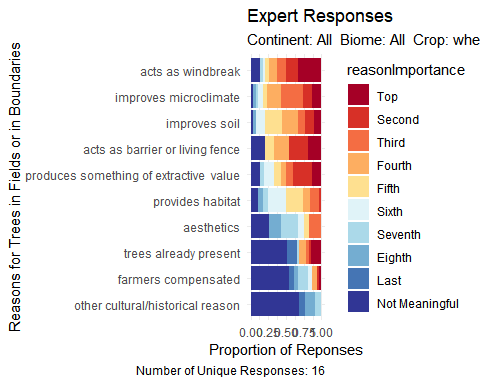
## # A tibble: 1,710 x 14  
## continent biome repTree name id institution countryExpertise  
## <chr> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 2 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 3 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 4 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 5 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 6 Africa Tropical & Subtro~ Faidhe~ Denn~ 4 ICRAF and ~ Many countries ~  
## 7 Africa Tropical & Subtro~ Alnus ~ Alai~ 5 Internatio~ Rwanda; Ethiopi~  
## 8 Africa Tropical & Subtro~ Alnus ~ Alai~ 5 Internatio~ Rwanda; Ethiopi~  
## 9 Africa Tropical & Subtro~ Alnus ~ Alai~ 5 Internatio~ Rwanda; Ethiopi~  
## 10 Africa Tropical & Subtro~ Alnus ~ Alai~ 5 Internatio~ Rwanda; Ethiopi~  
## # ... with 1,700 more rows, and 7 more variables: reasonImportance <fct>,  
## # reason <chr>, crop <chr>, b\_or\_f <chr>, policyPriority <fct>, policy <chr>,  
## # reasonValue <dbl>  
## [1] "acts as barrier or living fence"   
## [2] "produces something of extractive value"  
## [3] "Other"   
## [4] "improves soil"   
## [5] "improves microclimate"   
## [6] "acts as windbreak"   
## [7] "aesthetics"   
## [8] "trees already present"   
## [9] "provides habitat"   
## [10] "farmers compensated"   
## [11] "other cultural/historical reason"



1. Display plot for Wheat

crop\_histogram(crop="wheat",data=policy\_long,what2plot=what2plot,continent=continent,biome=biome)

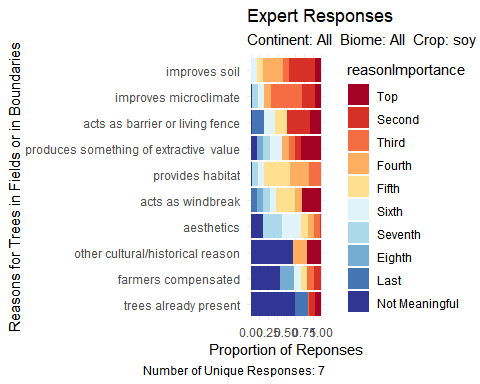
## # A tibble: 1,649 x 14  
## continent biome repTree name id institution countryExpertise  
## <chr> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Asia (includes Middle~ Temp~ Populu~ Huas~ 8 "Hebei Agr~ "China"   
## 2 Asia (includes Middle~ Temp~ Populu~ Amit~ 14 "Forest Re~ "India "   
## 3 Asia (includes Middle~ Temp~ Populu~ Amit~ 14 "Forest Re~ "India "   
## 4 Asia (includes Middle~ Temp~ Populu~ Amit~ 14 "Forest Re~ "India "   
## 5 Asia (includes Middle~ Temp~ Populu~ Amit~ 14 "Forest Re~ "India "   
## 6 Asia (includes Middle~ Temp~ Populu~ Amit~ 14 "Forest Re~ "India "   
## 7 Asia (includes Middle~ Temp~ Populu~ Amit~ 14 "Forest Re~ "India "   
## 8 Africa Temp~ Faidhe~ Tesf~ 16 "CIMMYT" "Ethiopia; Keny~  
## 9 Africa Temp~ Faidhe~ Tesf~ 16 "CIMMYT" "Ethiopia; Keny~  
## 10 Africa Temp~ Faidhe~ Tesf~ 16 "CIMMYT" "Ethiopia; Keny~  
## # ... with 1,639 more rows, and 7 more variables: reasonImportance <fct>,  
## # reason <chr>, crop <chr>, b\_or\_f <chr>, policyPriority <fct>, policy <chr>,  
## # reasonValue <dbl>  
## [1] "produces something of extractive value"  
## [2] "improves microclimate"   
## [3] "acts as barrier or living fence"   
## [4] "acts as windbreak"   
## [5] "improves soil"   
## [6] "farmers compensated"   
## [7] "trees already present"   
## [8] "aesthetics"   
## [9] "provides habitat"   
## [10] "Other"   
## [11] "other cultural/historical reason"



1. Display plot for Soy

crop\_histogram(crop="soy",data=policy\_long,what2plot=what2plot,continent=continent,biome=biome)

## # A tibble: 596 x 14  
## continent biome repTree name id institution countryExpertise  
## <chr> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Asia (includes Middle~ Temp~ "Popul~ "Hua~ 8 Hebei Agri~ "China"   
## 2 Africa Trop~ "Markh~ "Ath~ 15 CIFOR-ICRAF "Rwanda "   
## 3 Africa Trop~ "Markh~ "Ath~ 15 CIFOR-ICRAF "Rwanda "   
## 4 Africa Trop~ "Markh~ "Ath~ 15 CIFOR-ICRAF "Rwanda "   
## 5 Africa Trop~ "Markh~ "Ath~ 15 CIFOR-ICRAF "Rwanda "   
## 6 Africa Trop~ "Markh~ "Ath~ 15 CIFOR-ICRAF "Rwanda "   
## 7 Africa Trop~ "Markh~ "Ath~ 15 CIFOR-ICRAF "Rwanda "   
## 8 North/Central America~ Temp~ "sugar~ "And~ 21 University~ "Canada/other t~  
## 9 North/Central America~ Temp~ "sugar~ "And~ 21 University~ "Canada/other t~  
## 10 North/Central America~ Temp~ "sugar~ "And~ 21 University~ "Canada/other t~  
## # ... with 586 more rows, and 7 more variables: reasonImportance <fct>,  
## # reason <chr>, crop <chr>, b\_or\_f <chr>, policyPriority <fct>, policy <chr>,  
## # reasonValue <dbl>  
## [1] "produces something of extractive value"  
## [2] "trees already present"   
## [3] "improves soil"   
## [4] "acts as barrier or living fence"   
## [5] "acts as windbreak"   
## [6] "aesthetics"   
## [7] "farmers compensated"   
## [8] "improves microclimate"   
## [9] "provides habitat"   
## [10] "other cultural/historical reason"



# Questions and Next Steps

## Graphing options

1. At your request I can add other user inputs (for different combinations of biome, continent, policy recommendations/reasons for fields/reasons for boundaries)
2. I can also export plots from R as jpegs.
3. Colors, style, etc are mutable.
4. We can make other kinds of graphs: pie charts, side-by-side bar charts, and more.

## I have three data questions:

1. How do you want to handle the rare instances where experts typed in their own “reason” or “policy recommendation”

* Exclude these from quantitative analysis but mention them in discussion
* Choose the reason that is most closely connected to the reason they describe, but mention their nuance in the discussion

1. One expert (Pierluigi) filled out the questionnaire three times, each for a unique representative tree. How do you want to handle this?

* keep just one of Pierluigi’s responses, so that the “unique responses” are just “one per expert”
* expand all representative tree species columns for every expert, and treat them as having a unique (though duplicated) response for each tree
* do nothing, allow Pierluigi to have contributed three unique responses

1. Do you consider it ever meaningful to combine “reasons trees in boundaries” and “reasons trees in fields” for comparison between continent-biome pairs?

* I did so in this document just to demonstrate
* it’s also possible to keep these totally separate
* it’s slightly misleading to combine them, because only 8/9 of the options overlapped.