MCC DISCLOSURE REVIEW BOARD COVER SHEET

EVALUATION TITLE: MOROCCO – IRRIGATION BASELINE
DATA DE-IDENTIFICATION PACKAGE: BASELINE DATA
EVALUATOR NAME: JOSEPH GREEN
SUBMISSION DATE: DECEMBER 15, 2017
DRB REVIEW DATE:

1. **Introduction to Data Package**

(*Instructions: Include a paragraph summarizing each data package component included in the package. For example, if the package includes household, individual, and community level data sets, please include a paragraph summarizing each of these three components, including information on the content and timing of the data collection*.)

This data package includes the following components:

1. **Data Package Component 1: Complete Data**
   This survey is designed to establish a baseline (pre-project status) of the indicators used in monitoring and evaluating the Fruit Trees Project. It includes 8 questionnaires covering:
   - Professional Organizations
   - General data
   - Almond Trees
   - Other Cultures
   - Date-palm Trees
   - Olive Trees
   - Animal Production
   - Valuation Unit
   The study was conducted in three zones:
   - The small and medium hydraulic zone (PMH)
   - The Oasis zone, and
   - The Rain-Fed (Pluvial) zone. This zone has two different survey locations.
   The data is disaggregated into 46 files, grouped by questionnaire section and zone.

Complementary data collection efforts:
(*Instructions: Complementary data collection efforts are those efforts that complemented the data packages under review for de-identification, but do not necessarily require de-identification. The evaluator should list these data and provide a brief summary on how they connect to any data package components and affect the data package components' de-identification. For example, if the geospatial data for the project infrastructure is collected and will be publicly released, it should be listed in the complementary data collection efforts*.)

I am not aware of any complementary data collection efforts.

2. De-identification Data Package Folder Contents

(*Instructions: Please list the Data Package Component File Name, and then include the File Names of each of the corresponding required documents [Metadata, De-identification Worksheet, Informed Consent, Questionnaire, Other docs]. Only one de-identification worksheet per evaluation is requested unless discussed*

Data Package				
Component	**De-identification Worksheet**	**Informed Consent**	**Questionnaire**	**Other Documents**
Oasis_Organisations_Professionnelles.dta Pluvial1_Organisations_Professionnelles.dta PMH_Organisations_Professionnelles.dta			Fiche _Organisation_Professionnelle.doc	
Oasis_Donnees_Generales.dta Oasis_Parcellaire.dta Pluvial1_Donnees_Generales.dta Pluvial1_Parcellaire.dta Pluvial2_Donnees_Generales.dta Pluvial2_Parcellaire.dta PMH_Donnees_Generales.dta PMH_Parcellaire.dta			Fiche Donnees_Generales.doc	
Oasis_Amandier.dta Oasis_Variete_Amandier.dta Pluvial1_Amandier.dta Pluvial1_Variete_Amandier.dta Pluvial2_Amandier.dta Pluvial2_Variete_Amandier.dta PMH_Amandier.dta PMH_Variete_Amandier.dta	26 - Morocco - Irrigation Baseline - De-Identification Worksheet.docx		Fiche_Amandier.doc	File 1) 26 MAR_Irrigation Baseline - data info.xlsx File 2) 26 MAR_Irrigation Baseline - anonid_cross walk DO NOT PUBLISH.xlsx
Oasis_AutresCultures_Cout_Prod.dta Oasis_AutresCultures_Sup_Prod.dta Pluvial1_AutresCultures_Cout_Prod.dta Pluvial1_AutresCultures_Sup_Prod.dta Pluvial2_AutresCultures_Cout_Prod.dta Pluvial2_AutresCultures_Sup_Prod.dta PMH_AutresCultures_Cout_Prod.dta PMH_AutresCultures_Sup_Prod.dta			Fiche_Autres_Cultures.doc	
Oasis_Dattier.dta Oasis_Variete_Dattier.dta PMH_Dattier.dta PMH_Variete_Dattier.dta			Fiche_Dattier.doc	
Oasis_Olivier.dta Oasis_Variete_Olivier.dta Pluvial1_Olivier.dta Pluvial1_Variete_Olivier.dta Pluvial2_Olivier.dta Pluvial2_Variete_Olivier.dta PMH_Olivier.dta PMH_Variete_Olivier.dta			Fiche_Olivier.doc	
Oasis_Productions_Animales.dta Pluvial1_Productions_Animales.dta Pluvial2_Productions_Animales.dta PMH_Productions_Animales.dta			Fiche_Productions_animales.doc	
Oasis_Unite_Valorisation.dta Pluvial1_Unite_Valorisation.dta PMH_Unite_Valorisation.dta			Fiche_Unite_Valorisation.doc	

Country/Survey/Study Title:

Consultants should complete the following *De-Identification Worksheet* and include it in the delivery of public use, deidentified data.

The MCC Disclosure Review Board asks that the individuals responsible for de-identification prepare a response to the following questions in addition to this worksheet in preparation for the DRB review of the data:

1. **Understand De-Identification Process**: What was the general protocol/process followed for assessing risk and determining risk mitigation measures?
   RESPONSE: This survey is designed to establish a baseline (pre-project status) of the indicators used in monitoring and evaluating the Fruit Trees Project. There were nearly 10,000 variables across 46 data sets. The survey covered agricultural holdings, valuation units, and professional organizations. The survey contained the usual geo-political area names, and some organization and farmer names. But most sensitive information was in the form of loan amounts, agricultural holdings and sales, labor force numbers, and dates of transactions. The identity of the respondents can be masked by creating anonymous codes for the smaller geo-political areas, removing all respondent organization and farmer names, grouping loan amounts into ranges, aggregating dates to the year and month only, and recoding outlier values of agricultural holdings, sales, and number of workers into outlier medians.

2. **Understand Population Sizes**: What is the lowest geographic unit that can remain identified in the data based on risk assessment? What are the average population sizes by geographic unit?
   RESPONSE: The survey covered 4 zones, 21 provinces, and 58 perimeters. There is an average population of 462,000 in each province. Morocco has communes below the province level, but the communes are not identified in the data. The perimeters are below the level of the commune. Given that the communes can have populations as low as 5,000, identifying small perimeters within a commune puts respondents at risk of being re-identified. I kept provinces and zones fully identified by their names, and created anonymous codes for perimeter and questionnaire number.

Risk Level: (*Please rate High, Medium, Low):* Low

Informed Consent Statement: *Has the informed consent been submitted to MCC for documentation? Does the language in the informed consent statement allow for restricted use of complete data with personally identifiable information?*

IRB Clearance: *Does the IRB approval allow for release of restricted use of complete data with personally identifiable information?*

Issues		**Risk Analysis**		**Risk Mitigation**[1]	
		Instructions	*Response*	*Instructions*	*Response*
1.	Individuals or organizations that might have significant financial, legal, cultural, or other incentives to identify survey respondents	*List all potential users*	Government tax collection agencies that target agricultural holdings and transactions.		

---

[1] Any risk mitigation involving micro-aggregation, data swapping, adding noise or re-sampling should be documented as discussed in optional section below.

Country/Survey/Study Title:

	Issues	Risk Analysis		Risk Mitigation[1]	
		Instructions	*Response*	*Instructions*	*Response*
2.	Value or potential use of the data, if compromised. For example, some data might be used to capture delinquent tax payments, or could stigmatize the respondent	*List all uses*	The data might be used to collect taxes based on agricultural holdings and transactions.		
3.	Cost of re-identification, both in terms of costs to potential intruders to de-anonymize data and the cost to potential victims of identification	*Describe degree of difficulty for re-identification*	It would be very difficult to re-identify the data after all the de-identification measures I implemented The cost to potential victims would be negligible, since the most sensitive information on agriculture holdings and transactions were grouped into ranges or grouped at the top of their ranges.		
4.	Direct & indirect identifiers (names, addresses, geographic information, government-issued ID numbers, etc.)	*List the identifying items/variables.*	Zone, province, and perimeter names: zone province perim Organization and landholder names.	*List all identifying items/variables removed.*	Zone and province were kept in the public use data because they do not add significant risk of re-identification. I created anonymous values for perim and removed the perim names. I removed all names.
5.	GIS/GPS data	*List all GPS and/or GIS data.*	No GPS or GIS data.	*Describe process for de-identification. For example: introduce random errors into geographic data (GPS, GIS, etc.).*	N/A

Country/Survey/Study Title:

Issues		Risk Analysis		Risk Mitigation[1]	
		Instructions	*Response*	*Instructions*	*Response*
				Displace urban points 0-2 km, rural points 0-5 km, and additional 1% of rural points 0-10 km[2].	
6.	Outliers in data that allow identification of respondent	*List the identifying items/variables:*	Too many to list here. See Appendix A for details.	*Describe top/bottom coding: set upper & lower bounds to remove outliers for continuous. Specify: are values set to missing, the median, or other?* *For large categories/datasets, the OMB suggests top coding at least the highest .5%; for smaller categories/datasets, top code the highest 3-5%. The same principles apply to bottom coding.[3]*	See Appendix A.
				Describe any variables that require collapse and describe construction of new variable	See Appendix A.
				Describe any global re-coding to group observations into categories (e.g., age 0-5, 5-10, 65+, etc.). Ensure that the categories are neither too broad nor too narrow.	See Appendix A.
7.	Unique and rare observations, as well as highly visible variable combinations (for example: individuals with high incomes, ages, or unique combinations, such as 17-year old widowers or contextually unusual racial/ethnic backgrounds)	*List the identifying items/variables:*	Too many to list here. See Appendix A for details.	*For each identified rare data, describe the local suppression techniques employed to remove unique and rare data. Specify: are values set to missing, the median, or other?* *Example: Replace observations within a unique/rare record with missing values (e.g., replacing 'marital status*	See Appendix A.

---

[2] ICF International, Demographic & Health Surveys
[3] Office of Management and Budget, Checklist on Disclosure Potential of Proposed Data Releases (current link)

Country/Survey/Study Title:

Issues	Risk Analysis		Risk Mitigation[1]	
	Instructions	*Response*	*Instructions*	*Response*
			= widow; age = 17' with 'marital status =. ; age = 17')	
			Describe any global re-coding to group observations into categories (e.g., age 0-5, 5-10, 65+, etc.). *Ensure that the categories are neither too broad nor too narrow.*	See Appendix A.
8. Availability of 'key' data that can be used to identify respondents. This includes other datasets or archives with information that can be used to re-identify individuals in the dataset.	*List all potential existing data*	Questionnaire number, phone numbers, addresses	*Describe how to mitigate link to existing data that enables re-identification*	I anonymized the original values for the questionnaire ID by replacing them with randomly generated values. I removed all addresses and phone numbers.
9. Additional requirements by the Institutional Review Board(s) for data de-identification	*List all requirements*		*List responses to requirements*	

Optional Section – Complete as necessary

Risk Mitigation – Data Perturbation

1. For sensitive datasets, consider 'micro-aggregation' – average observations within *k* records to ensure that at least *k* units within a dataset are identical. For example, for the variable *age* and if *k*=4, take 4 observations, average their age, and replace each record's *age* with the average.

   Describe any micro-aggregation employed, and list the affected variables:

None.

2. For sensitive data, consider 'data-swapping' – swapping values between records. This can be targeted (e.g., only unique records) or random; however, care must be taken not to alter the data too significantly.

   Describe any data-swapping employed, and list the affected variables:

None.

3. Add noise or resampling for selected variables. *Note that extensive data perturbation (i.e., perturbation of many variables) is not recommended as these processes can degrade data significantly*.

   If noise was added or resampling occurred, please list the affected variables, why this procedure was deemed necessary and efforts taken to avoid degrading the data.

None.

Appendix A: Data manipulation summary

Oasis

Oasis_Amandier
Created anonymous ID codes: code perim nfiche

Oasis_AutresCultures_Cout_Prod
Created anonymous ID codes: code perim nfiche

Oasis_AutresCultures_Sup_Prod
Created anonymous ID codes: code perim nfiche

Recoded outliers to the median of the outliers: AC1_SUPCULTURE

Oasis_Dattier
Created anonymous ID codes: code perim nfiche

Oasis_Donnees_Generales
Created anonymous ID codes: code perim nfiche DS1_COMMUNE DS1_DOUAR

Removed variables with PII that either cannot be de-identified, or are not useful for analysis:
DS1_ENQUETEUR DS1_SUPERVISEUR

Round specific dates to the year + month: DS1_DATENQ DS11_COMP1DEB DS11_COMP1FIN

Group infrequent categories: DS2_SITUATION

Recoded outliers to the median of the outliers: DS2_AGE

Remove values from variables with very few responses: DS9_FEMNBR DS9_FEMCHARGE

Grouped values into ranges: DS11_COMP1MONTANT DS11_COMP1ANNUITE

Oasis_Olivier
Created anonymous ID codes: code perim nfiche

Oasis_Organisations_Professionnelles
Created anonymous ID codes: code perim nfiche

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis:
nbrehom nbrefem acthom actfem nbrvehicule SECRPERM2 nbrfemca dettescumul

Rounded specific dates to month + year: dateenquet dateag

Grouped values into ranges: datecrea nbradher nbractuel superficie montantcotis capitalsocial capitallibere
resultatnet

Recoded outlier values to outlier medians: fraisadhe

Aggregated gender-specific # of workers to total # of workers: tot nbrsans

Oasis_Parcellaire
Created anonymous ID codes: code perim nfiche

Grouped values into ranges: DS10_SUPERF DS10_SUPJAC

Recoded outliers to outlier medians: DS10_PIEDOLIVIER DS10_PIEDDATIER DS10_PIEDAMAND
DS10_PIEDAUTR

Oasis_Productions_Animales
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: AN11_OVINMAL AN11_OVINFEM AN11_OVINJEUN
AN24_BOVGMOFM AN24_BOVGMOFF AN24_BOVGMOSM AN24_BOVGMOSF AN24_BOVAMOFM
AN24_BOVAMOFF AN24_BOVAMOSM AN24_BOVAMOSF AN24_BOVEMOFM AN24_BOVEMOFF
AN24_BOVEMOSM AN24_BOVEMOSF AN24_BOVSMM AN24_BOVSMF AN24_CAMGMOFM
AN24_CAMAMOFM AN24_CAMEMOFM AN31_DEBEFFBOV AN31_DEBEFFOV AN31_DEBEFFCAP
AN31_DEBEFFAVI AN31_DEBEFFCUN AN31_DEBEFFAPI AN31_DEBVALBOV AN31_DEBVALOV
AN31_DEBVALCAP AN31_DEBVALEQ AN31_DEBVALAVI AN31_DEBVALCUN AN31_DEBVALAPI
AN31_VNTEFFBOV AN31_VNTEFFOV AN31_VNTEFFCAP AN31_VNTVALBOV AN31_VNTVALOV
AN31_VNTVALCAP AN31_VNTVALEQ AN31_ACHEFFBOV AN31_ACHEFFOV AN31_ACHVALBOV
AN31_ACHVALOV AN31_ACHVALCAP AN31_NASEFFBOV AN31_NASEFFOV AN31_NASEFFCAP
AN31_NASEFFAVI AN31_MOREFFOV AN31_FINEFFBOV AN31_FINEFFOV AN31_FINEFFCAP
AN31_FINEFFEQ AN31_FINEFFAVI AN31_FINEFFCUN AN31_FINEFFAPI AN31_FINVALBOV
AN31_FINVALOV AN31_FINVALCAP AN31_FINVALEQ AN31_FINVALAVI AN31_FINVALCUN
AN31_FINVALAPI AN11_CAPMAL AN11_CAPFEM AN11_CAPJEUN AN12_VOLAILLE AN12_LAPIN
AN12_RUCHTRA AN12_RUCHMOD AN31_DEBEFFEQ

Recoded unique #s to "1+": AN24_CAMAMOFF AN24_PETAMOFF AN24_PETEMOFF
AN24_APIGMOFM AN24_APIGMOSM AN24_APIAMOFM AN24_APIEMOFM AN24_APIEMOSM
AN24_APISMM AN31_VNTEFFCUN AN31_VNTVALAVI AN31_VNTVALCUN AN31_ACHEFFCAP
AN31_ACHVALEQ AN31_ACHVALAVI AN31_ACHEFFAVI AN31_MOREFFBOV

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis:
AN31_VNTEFFAVI AN31_NASEFFEQ AN31_MOREFFAPI AN31_DONEFFBOV

Oasis_Unite_Valorisation
Created anonymous ID codes: code perim nfiche

Removed variables with PII that either cannot be de-identified, or are not useful for analysis: enqueteur
UV1_NOM_UNITE UV1_ADRESSE UV1_RESPONSABLE UV1_TEL UV8_DEMFOURN
UV9_SALONFOURN UV9_FOIRFOUR

Aggregated specific date to the month + year: datenq UV2_DATASSOCIAT

Recoded infrequent values to missing: UV1_STATUT

Aggregated specific date to year ranges: UV1_DATCREAT

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis: UV3_OUVFEM

Grouped values into ranges: UV3_JOURHOM UV3_JOURFEM UV61_DATCAP UV61_DATTRAIT UV62_OLIVPREST UV62_DATPREST UV71_DATTES UV72_DATPROD UV72_DATQTE1 UV72_DATVAL1 UV61_OLIVRECET UV61_DATRECET

Recoded unique #s to "1+": UV71_OLIVES UV71_AMANDES UV72_HUILORDPROD UV72_HUILORDPART UV72_HUILORDQTE1 UV72_HUILORDVAL1 UV72_PATPROD UV72_PATPART UV72_PATQTE1 UV72_PATVAL1 UV72_CONFPROD UV72_CONFQTE1 UV72_CONFVAL1 UV72_MIELPROD UV72_MIELQTE1 UV72_MIELVAL1 UV72_JUSPROD UV72_JUSQTE1 UV72_JUSVAL1

Oasis_Variete_Amandier
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: AM1_AMDSUR

Oasis_Variete_Dattier
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: DT1_DATSUR DT1_DATHORS

Oasis_Variete_Olivier
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: OL1_PIEDSUR OL1_PIEDHORS

Pluvial1

Pluvial1_Amandier
Created anonymous ID codes: code perim nfiche

Pluvial1_AutresCultures_Cout_Prod
Created anonymous ID codes: code perim nfiche

Pluvial1_AutresCultures_Sup_Prod
Created anonymous ID codes: code perim nfiche

Recoded outlier areas to the median of the outliers: AC1_SUPCULTURE

Pluvial1_Donnees_Generales
Created anonymous ID codes: code perim nfiche DS1_COMMUNE DS1_DOUAR

Removed variables with PII that either cannot be de-identified, or are not useful for analysis: DS1_ENQUETEUR DS1_SUPERVISEUR

Aggregated specific date to month + year: DS1_DATENQ

Recoded infrequent marriage status "other": DS2_SITUATION

Recoded outlier values to the median of the outliers: DS2_AGE DS9_MASCHARGE

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis: DS9_FEMNBR DS9_FEMCHARGE

Grouped values into ranges: DS11_COMP1MONTANT DS11_INVT1MONTANT DS11_INVT1ANNUITE DS11_INVT2MONTANT

Replaced specific loan dates with rounded year values: DS11_COMP1DEB DS11_COMP1FIN DS11_INVT1DEB DS11_INVT1FIN DS11_INVT2DEB DS11_INVT2FIN

Pluvial1_Olivier
Created anonymous ID codes: code perim nfiche

Pluvial1_Organisations_Professionnelles
Created anonymous ID codes: code perim nfiche

Removed variables with PII that either cannot be de-identified, or are not useful for analysis: enqueteur nomopa adresseopa president tel

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis: nbrtechh

Aggregate specific date variable to month + year: dateenquet dateag

Aggregate date of formation values to year, and group values together: datecrea

Grouped values into ranges: percouvert datecrea nbradher nbrehom nbrefem fraisadhe nbractuel acthom actfem conseiladm conseilhom conseilfem superficie totpersonnel tothom montantcotis capitalsocial capitallibere resultatnet

Pluvial1_Parcellaire
Created anonymous ID codes: code perim nfiche

Grouped values into ranges: DS10_SUPERF DS10_SUPOLIV DS10_SUPCER DS10_SUPLEG DS10_SUPFOUR

Recoded outliers to outlier medians: DS10_PIEDOLIVIER DS10_PIEDAMAND DS10_PIEDAUTR

Pluvial1_Productions_Animales
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: AN11_BOVAMEMAL AN11_BOVAMEFEM AN11_BOVAMEJEUN AN11_BOVLOCMAL AN11_BOVLOCFEM AN11_BOVLOCJEUN AN11_OVINMAL AN11_OVINFEM

AN11_OVINJEUN AN11_CAPMAL AN11_CAPFEM AN11_CAPJEUN AN11_EQDMAL AN11_EQDFEM
AN11_EQDJEUN AN12_VOLAILLE AN12_LAPIN AN12_RUCHTRA AN12_RUCHMOD
AN21_BOVQTECON AN21_BOVVALCON AN21_BOVQTEVRT AN21_BOVVALVRT
AN21_BOVQTEGROS AN21_BOVVALGROS AN21_OVQTECON AN21_OVVALCON
AN21_OVQTEVRT AN21_OVVALVRT AN21_OVQTEGROS AN21_OVVALGROS AN21_CAPQTECON
AN21_CAPVALCON AN21_CAPQTEGROS AN21_CAPVALGROS AN21_EQQTECON
AN21_EQVALCON AN21_EQQTEGROS AN21_EQVALGROS AN21_AVIQTECON AN21_AVIVALCON
AN21_BOVQTECON_Norm AN21_BOVVALCON_Norm AN21_BOVQTEVRT_Norm
AN21_BOVVALVRT_Norm AN21_BOVQTEGROS_Norm AN21_BOVVALGROS_Norm
AN21_OVQTECON_Norm AN21_OVVALCON_Norm AN21_OVQTEVRT_Norm
AN21_OVVALVRT_Norm AN21_OVQTEGROS_Norm AN21_OVVALGROS_Norm
AN21_CAPQTECON_Norm AN21_CAPVALCON_Norm AN21_CAPQTEGROS_Norm
AN21_CAPVALGROS_Norm AN21_EQQTECON_Norm AN21_EQVALCON_Norm
AN21_EQQTEGROS_Norm AN21_EQVALGROS_Norm AN21_AVIQTECON_Norm
AN21_AVIVALCON_Norm AN24_BOVGMOFM AN24_BOVGMOFF AN24_BOVGMOSM
AN24_BOVGMOSF AN24_BOVAMOFM AN24_BOVAMOFF AN24_BOVAMOSM AN24_BOVAMOSF
AN24_BOVEMOFM AN24_BOVEMOFF AN24_BOVEMOSM AN24_BOVEMOSF AN24_BOVSMF
AN24_CAMGMOFM AN24_CAMGMOFF AN24_CAMGMOSM AN24_CAMGMOSF
AN24_CAMAMOFM AN24_CAMAMOFF AN24_CAMAMOSM AN24_CAMAMOSF
AN24_CAMEMOFM AN24_CAMEMOFF AN24_CAMEMOSM AN24_CAMSMM AN24_CAMSMF
AN24_PETAMOFM AN24_PETAMOFF AN24_PETEMOFF AN31_DEBEFFBOV AN31_DEBEFFOV
AN31_DEBEFFCAP AN31_DEBEFFEQ AN31_DEBEFFAVI AN31_DEBEFFCUN AN31_DEBEFFAPI
AN31_DEBVALBOV AN31_DEBVALOV AN31_DEBVALCAP AN31_DEBVALEQ AN31_DEBVALAVI
AN31_DEBVALCUN AN31_DEBVALAPI AN31_VNTEFFBOV AN31_VNTEFFOV AN31_VNTEFFCAP
AN31_VNTEFFAVI AN31_VNTVALBOV AN31_VNTVALOV AN31_VNTVALCAP AN31_VNTVALEQ
AN31_VNTVALAVI AN31_ACHEFFBOV AN31_ACHEFFOV AN31_ACHEFFCAP AN31_ACHEFFEQ
AN31_ACHEFFAVI AN31_ACHVALBOV AN31_NASEFFBOV AN31_NASEFFOV AN31_NASEFFCAP
AN31_NASEFFAVI AN31_NASEFFCUN AN31_ACHVALOV AN31_ACHVALCAP AN31_ACHVALEQ
AN31_ACHVALAVI AN31_MOREFFOV AN31_MOREFFCAP AN31_MOREFFAVI AN31_FINEFFBOV
AN31_FINEFFOV AN31_FINEFFCAP AN31_FINEFFEQ AN31_FINEFFAVI AN31_FINEFFCUN
AN31_FINEFFAPI AN31_FINVALBOV AN31_FINVALOV AN31_FINVALCAP AN31_FINVALEQ
AN31_FINVALAVI AN31_FINVALCUN AN31_FINVALAPI AN24_BOVSMM AN24_PETSMM

Recoded unique #s to "1+": AN24_PETGMOFF AN24_PETGMOSF AN24_PETEMOFM AN31_VNTEFFEQ
AN31_NASEFFEQ AN31_MOREFFBOV AN31_DONEFFBOV

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis:
AN21_EQQTEVRT AN21_EQVALVRT AN21_EQQTEVRT_Norm AN21_EQVALVRT_Norm
AN21_CUNQTECON AN21_CUNVALCON AN21_CUNQTECON_Norm AN21_CUNVALCON_Norm
AN21_APIQTECON AN21_APIVALCON AN21_APIQTECON_Norm AN21_APIVALCON_Norm
AN24_PETGMOSM AN24_PETAMOSF AN24_PETEMOSM AN24_PETEMOSF AN24_APIGMOFM
AN24_APIAMOFM AN24_APIAMOFF AN24_APIAMOSM AN24_APIEMOFM AN24_APISMM
AN24_APISMF AN31_VNTEFFCUN AN31_VNTVALCUN AN31_ACHEFFAPI AN31_ACHVALAPI
AN31_NASEFFAPI AN31_MOREFFCUN AN31_MOREFFAPI AN31_PRIEFFBOV AN31_PRISEFFOV
AN31_PRIVALBOV AN31_PRIVALOV AN31_DONVALBOV

Pluvial1_UnitÇ_Valorisation
Created anonymous ID codes: code perim nfiche

Removed variables with PII that either cannot be de-identified, or are not useful for analysis: enqueteur UV1_NOM_UNITE UV1_ADRESSE UV1_RESPONSABLE UV1_TEL

Aggregated specific date to the month + year: datenq

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis: UV1_GENRE UV1_SEXE UV2_ASSOCIATION UV41_ATRCADMAS UV72_OLIVNPROD UV72_OLIVNPROD UV72_OLIVVPROD UV72_HUILORDQTE2 UV72_HUILORDVAL2 UV72_HUILEXTPROD UV72_HUILEXTPART UV72_HUILEXTQTE1 UV72_HUILEXTVAL1 UV72_JUSPROD

Recoded infrequent values to "other": UV1_STATUT

Aggregated specific date to year ranges: UV1_DATCREAT

Recoded outliers to outlier medians: UV3_JOURHOM UV3_JOURFEM UV41_ATRATRMAS UV5_INDTOT UV5_INDETAT UV5_ARTTOT UV5_ARTETAT UV5_ATR1TOT UV5_ATR1ETAT UV61_OLIVCAP UV61_OLIVTRAIT UV61_OLIVQTE UV61_OLIVRECET UV62_OLIVPREST UV71_OLIVES UV72_HUILORDPROD UV72_HUILORDPART UV72_HUILORDQTE1 UV72_HUILORDVAL1 UV72_HUILVRGPROD UV72_HUILVRGQTE1 UV72_HUILVRGVAL1 UV10_MARGQTE UV10_TOURQTE UV10_ATR1QTE UV5_ATR2TOT UV5_ATR2ETAT UV72_HUILVRGPART

Recoded unique #s to "1+": UV5_ATR3TOT UV5_ATR3ETAT

Pluvial1_Variete_Amandier
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: AM1_AMDSUR

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis: AM1_AMDHORS

Pluvial1_Variete_Olivier
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: OL1_PIEDSUR OL1_PIEDHORS

Pluvial2

Pluvial2_Amandier
Created anonymous ID codes: code perim nfiche

Pluvial2_AutresCultures_Cout_Prod
Created anonymous ID codes: code perim nfiche

Pluvial2_AutresCultures_Sup_Prod
Created anonymous ID codes: code perim nfiche

Pluvial2_Donnees_Generales
Created anonymous ID codes: code perim nfiche DS1_COMMUNE DS1_DOUAR

Removed variables with PII that either cannot be de-identified, or are not useful for analysis:
DS1_ENQUETEUR DS1_SUPERVISEUR

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis:
DS2_SEXE

Aggregate specific date to month + year: DS1_DATEAN

Recoded infrequent marriage status or "other": DS2_SITUATION

Recoded outlier values to the median of the outliers: DS2_AGE

Pluvial2_Olivier
Created anonymous ID codes: code perim nfiche

Pluvial2_Parcellaire
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: DS10_PIEDOLIVIER DS10_PIEDAMAND

Pluvial2_Productions_Animales
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: AN11_BOVLOCFEM AN11_BOVLOCJEUN AN11_OVINMAL
AN11_OVINFEM AN11_OVINJEUN AN11_CAPFEM AN11_CAPJEUN AN11_EQDFEM
AN24_BOVGMOFM AN24_BOVGMOFF AN24_BOVAMOFM AN24_BOVAMOFF AN24_BOVEMOFF
AN24_CAMGMOFM AN24_CAMGMOFF AN24_CAMAMOFM AN24_CAMAMOFF
AN24_CAMEMOFM AN31_DEBEFFBOV AN31_DEBEFFOV AN31_DEBEFFCAP AN31_DEBEFFEQ
AN31_DEBVALBOV AN31_DEBVALOV AN31_DEBVALCAP AN31_DEBVALEQ AN31_DEBVALAVI
AN31_VNTEFFBOV AN31_VNTEFFOV AN31_VNTVALBOV AN31_VNTVALOV AN31_ACHEFFOV
AN31_ACHVALOV AN31_AUTVALAVI AN31_NASEFFBOV AN31_NASEFFOV AN31_MOREFFOV
AN31_FINEFFBOV AN31_FINEFFOV AN31_FINEFFCAP AN31_FINEFFEQ AN31_FINVALBOV
AN31_FINVALOV AN31_FINVALCAP AN31_FINVALEQ AN31_FINVALAVI AN24_BOVEMOFM
AN31_FINEFFAVI AN24_CAMEMOFF AN31_DEBEFFAVI AN31_FINEFFEQ

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis:
AN11_BOVLOCMAL AN11_CAPMAL AN12_RUCHTRA AN12_RUCHMOD AN21_CAPQTECON
AN21_CAPVALCON AN22_CAP1QTECON AN22_CAP1VALCON AN22_CAP1QTEGROS
AN22_CAP1VALGROS AN21_CAPQTECON_Norm AN21_CAPVALCON_Norm
AN22_CAP1QTECON_Norm AN22_CAP1VALCON_Norm AN22_CAP1QTEGROS_Norm
AN22_CAP1VALGROS_Norm AN24_BOVGMOSM AN24_BOVAMOSM AN24_BOVEMOSM
AN31_VNTEFFCAP AN31_VNTEFFAVI AN31_VNTVALCAP AN31_VNTVALAVI AN31_ACHEFFBOV
AN31_ACHEFFEQ AN31_ACHEFFAVI AN31_ACHVALBOV AN31_ACHVALEQ AN31_ACHVALAVI
AN31_AUTEFFCAP AN31_AUTVALCAP AN31_NASEFFCAP AN31_NASEFFEQ AN31_NASEFFAVI
AN31_MOREFFBOV AN31_MOREFFCAP AN31_MOREFFAVI

Recoded unique #s to "1+": AN31_AUTEFFAVI

Pluvial2_Variete_Amandier
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: AM1_AMDSUR

Pluvial2_Variete_Olivier
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: OL1_PIEDSUR

PMH

PMH_Amandier
Created anonymous ID codes: code perim nfiche

PMH_AutresCultures_Cout_Prod
Created anonymous ID codes: code perim nfiche

PMH_AutresCultures_Sup_Prod
Created anonymous ID codes: code perim nfiche

Recode outlier areas to the median of the outliers: AC1_SUPCULTURE

PMH_Dattier
Created anonymous ID codes: code perim nfiche

PMH_Donnees_Generales
Created anonymous ID codes: code perim nfiche DS1_COMMUNE DS1_DOUAR

Removed variables with PII that either cannot be de-identified, or are not useful for analysis: DS1_ENQUETEUR DS1_SUPERVISEUR

Aggregate specific date to month + year: DS1_DATEAN

Recoded infrequent marriage status to "other": DS2_SITUATION

Recoded outlier values to the median of the outliers: DS2_AGE DS2_EXPERIENCE DS9_MASCHARGE DS11_COMP1MONTANT DS11_COMP1ANNUITE DS11_INVT1MONTANT DS11_INVT1ANNUITE

Removed values for variables with very few responses: DS9_FEMNBR DS9_FEMCHARGE

Rounded specific date of loan to year: DS11_COMP1DEB DS11_COMP1FIN DS11_INVT1DEB DS11_INVT1FIN

PMH_Olivier
Created anonymous ID codes: code perim nfiche

PMH_Organisations_Professionnelles
Created anonymous ID codes: code perim nfiche

Removed variables with PII that either cannot be de-identified, or are not useful for analysis: enqueteur nomopa adresseopa president tel

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis: conseilfem

Aggregated date to the year: dateenquet

Grouped values into ranges: datecrea nbradher nbrehom nbrefem nbractuel acthom actfem conseiladm conseilhom nbrsansh tauxag nbrmasag nbrfemag membreca reunionca fraisadhe superficie REUNIONCA09 montantcotis capitalsocial capitallibere resultatnet nbrvehicule dettescumul

Aggregated specific date to month + year: dateag

PMH_Parcellaire
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: DS10_SUPERF DS10_SUPCER DS10_PIEDOLIVIER DS10_PIEDDATIER DS10_PIEDAMAND DS10_SUPOLIV DS10_SUPJAC DS10_PIEDAUTR

PMH_Productions_Animales
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: AN11_BOVAMEMAL AN11_BOVAMEFEM AN11_BOVAMEJEUN AN11_BOVLOCMAL AN11_BOVLOCFEM AN11_OVINMAL AN11_OVINFEM AN11_OVINJEUN AN11_CAPMAL AN11_CAPFEM AN11_CAPJEUN AN11_EQDMAL AN12_RUCHTRA AN12_RUCHMOD AN21_BOVQTEVRT AN21_CAPQTECON AN21_CAPVALCON AN21_OVQTEVRT AN21_OVVALVRT AN21_CAPQTEGROS AN21_CAPVALGROS AN21_EQQTECON AN21_EQVALCON AN21_EQQTEGROS AN21_EQVALGROS AN21_AVIQTECON AN21_AVIVALCON AN21_CUNQTECON AN21_CUNVALCON AN21_APIQTECON AN21_APIVALCON AN21_BOVQTEVRT_Norm AN21_OVQTEVRT_Norm AN21_OVVALVRT_Norm AN21_CAPQTECON_Norm AN21_CAPVALCON_Norm AN21_CAPQTEGROS_Norm AN21_CAPVALGROS_Norm AN21_EQQTECON_Norm AN21_EQVALCON_Norm AN21_EQQTEGROS_Norm AN21_EQVALGROS_Norm AN21_AVIQTECON_Norm AN21_AVIVALCON_Norm AN21_CUNQTECON_Norm AN21_CUNVALCON_Norm AN21_APIQTECON_Norm AN21_APIVALCON_Norm AN24_BOVGMOFM AN24_BOVGMOFF AN24_BOVGMOSM AN24_BOVGMOSF AN24_BOVAMOFM AN24_BOVAMOFF AN24_BOVAMOSM AN24_BOVAMOSF AN24_BOVEMOFM AN24_BOVEMOFF AN24_BOVEMOSM AN24_BOVEMOSF AN24_BOVSMM AN24_BOVSMF AN24_CAMGMOFM AN24_CAMGMOFF AN24_CAMAMOFM AN24_CAMAMOFF AN24_CAMEMOFM AN24_CAMEMOFF AN24_PETGMOFF AN24_PETAMOFM AN24_PETAMOFF AN24_PETEMOFF AN24_PETSMF AN31_DEBEFFBOV AN31_DEBEFFOV AN31_DEBEFFCAP AN31_DEBEFFEQ AN31_DEBEFFAVI AN31_DEBEFFCUN AN31_DEBEFFAPI AN31_DEBVALBOV AN31_DEBVALOV AN31_DEBVALCAP AN31_DEBVALEQ AN31_DEBVALAVI AN31_DEBVALCUN AN31_DEBVALAPI AN31_VNTEFFBOV AN31_VNTEFFOV AN31_VNTEFFCAP AN31_VNTEFFAVI AN31_VNTEFFCUN AN31_VNTVALBOV AN31_VNTVALOV AN31_VNTVALCAP AN31_VNTVALAVI AN31_VNTVALCUN AN31_ACHEFFBOV AN31_ACHEFFOV AN31_ACHEFFCAP

AN31_ACHEFFAVI AN31_ACHEFFCUN AN31_ACHEFFAPI AN31_ACHVALBOV AN31_ACHVALOV
AN31_ACHVALCAP AN31_ACHVALAVI AN31_ACHVALCUN AN31_ACHVALAPI
AN31_NASEFFBOV AN31_NASEFFOV AN31_NASEFFCAP AN31_NASEFFEQ AN31_NASEFFAVI
AN31_NASEFFCUN AN31_MOREFFBOV AN31_MOREFFOV AN31_MOREFFCAP AN31_MOREFFAVI
AN31_MOREFFCUN AN31_MOREFFAPI AN31_FINEFFBOV AN31_FINEFFOV AN31_FINEFFCAP
AN31_FINEFFEQ AN31_FINEFFAVI AN31_FINEFFCUN AN31_FINEFFAPI AN31_FINVALBOV
AN31_FINVALOV AN31_FINVALCAP AN31_FINVALEQ AN31_FINVALAVI AN31_FINVALCUN
AN31_FINVALAPI AN31_PRIEFFBOV AN12_VOLAILLE AN12_LAPIN AN21_BOVQTECON
AN21_BOVVALCON AN21_BOVVALVRT AN21_BOVQTEGROS AN21_BOVVALGROS
AN21_OVQTECON AN21_OVVALCON AN21_OVQTEGROS AN21_OVVALGROS
AN21_BOVQTECON_Norm AN21_BOVVALCON_Norm AN21_BOVVALVRT_Norm
AN21_BOVQTEGROS_Norm AN21_BOVVALGROS_Norm AN21_OVQTECON_Norm
AN21_OVVALCON_Norm AN21_OVQTEGROS_Norm AN21_OVVALGROS_Norm

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis:
AN21_CUNQTEVRT AN21_CUNVALVRT AN21_CUNQTEVRT_Norm AN21_CUNVALVRT_Norm
AN24_CAMGMOSF AN24_CAMAMOSF AN24_CAMEMOSF AN24_PETGMOFM AN24_PETAMOSM
AN24_PETAMOSF AN24_PETEMOFM AN24_PETEMOSM AN24_PETEMOSF AN24_PETSMM
AN24_APIAMOFM AN24_APIEMOFM AN24_APIEMOFF AN24_APISMM AN24_APISMF
AN31_NASEFFAPI AN31_MOREFFEQ AN31_PRIVALBOV AN31_DONEFFOV AN31_DONEFFCAP
AN31_DONVALBOV AN31_DONVALOV AN31_DONVALCAP

PMH_Unite_Valorisation
Created anonymous ID codes: code perim nfiche

Removed variables with PII that either cannot be de-identified, or are not useful for analysis: enqueteur
UV1_NOM_UNITE UV1_ADRESSE UV1_RESPONSABLE UV1_TEL

Aggregated specific date to the month + year: datenq

Recoded infrequent values to missing: UV1_GENRE

Removed values for variables with PII that either cannot be de-identified, or are not useful for analysis:
UV1_SEXE UV3_CADREHOM UV3_TECHOM UV3_JOURFEM UV5_MACEXIST UV5_MACTOT
UV5_MACETAT UV5_LABTOT UV5_LABETAT UV5_ORDEXIST UV5_ORDTOT UV5_ORDETAT
UV72_OLIVNPROD UV72_OLIVNQTE1 UV72_OLIVNPRIX1 UV72_OLIVNVAL1 UV72_OLIVVPROD
UV72_OLIVVPART UV72_OLIVVQTE1 UV72_OLIVVPRIX1 UV72_OLIVVVAL1 UV9_ATR3FOURN

Recoded infrequent values to "other": UV1_STATUT

Aggregated specific date to year ranges: UV1_DATCREAT

Recoded outliers to outlier medians: UV3_TOTHOM UV3_OUVSANSHOM UV3_JOURHOM
UV5_INDTOT UV5_INDETAT UV5_TRANSTOT UV5_TRANSETAT UV61_OLIVCAP
UV61_OLIVDURE UV61_OLIVTRAIT UV61_OLIVQTE UV61_OLIVRECET UV62_OLIVPREST
UV71_OLIVES UV72_HUILORDPROD UV72_HUILORDPART UV72_HUILORDQTE1
UV72_HUILORDVAL1 UV72_HUILVRGPROD UV72_HUILVRGPART UV72_HUILVRGQTE1
UV72_HUILVRGVAL1 UV10_MARGQTE UV10_TOURQTE UV72_HUILEXTPROD
UV72_HUILEXTPART UV72_HUILEXTQTE1 UV72_HUILEXTVAL1

Recoded unique #s to "1+": UV3_OUVHOM

<u>PMH_Variete_Amandier</u>
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: AM1_AMDSUR

<u>PMH_Variete_Dattier</u>
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: DT1_DATSUR

<u>PMH_Variete_Olivier</u>
Created anonymous ID codes: code perim nfiche

Recoded outliers to outlier medians: OL1_PIEDSUR OL1_PIEDHORS