

Systems and Computer Engineering Department (SYSC)

Ottawa Collision Data Report:

Brief Clean and Homologation Transformations

Enrique Reverón Yanes

January 6th, 2019

**Ottawa Collision Data Report**

**About this Document**

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Remarks** |
| V1.0 | January 6th 2018 | Enrique Reverón | First draft version |

**Table of Contents**

[Introduction 4](#_Toc534563326)

[1. Brief Homologation and Clean Transformations 5](#_Toc534563327)

[References 5](#_Toc534563328)

# Introduction

This document includes a brief information about the documents *Ottawa Collision Report: Homologation Transformations* and *Ottawa Collision Report: Clean Transformations*.

# Brief Homologation and Clean Transformations Table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Feature Name** | **Feature Value** | **Counter of Observations** | **Percentage of Observations** | **Datasets** | **Homologation / Clean Transformation** | **Comments / Suggestions** |
| **XCOORD / YCOORD** | "" (Blank) | 28453 | 7.5000 | 2009 | The 2009 dataset do not have the coordinates values |  |
| 0 | 716 | 0.2000 | 2014 - 2015 | Removed the observations related |  |
| Very High / Low Values | Looks like some coordinates are outside of Ottawa Map (to be validated), my idea to handle this is just removed the observations related | | | |  |
| **ACCIDENT\_DATE** | The 2004-2012 datasets include the observations of first day of the next year (for example the 2004 dataset includes 2005-01-01 observations), I already remove those values on each dataset (keep only the ones that belongs to each year). | | | | |  |
| **DRIVER\_SEX** | "" (Blank) | 17 | 0.0045 | 2015-2016 | Those observations could he homologated as "00 - Unknown"? |  |
| 00 - Unknown | 21411 | 5.6000 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **DRIVER\_AGE** | I understand that the value is the driver age at the moment of the collision, that's right? | | | | |  |
| < 0 | 2 | 0.0006 | 2015 | Removed the observations related |  |
| 0 | 21180 | 5.5000 | 2004-2017 | Removed the observations related |  |
| >90 | 290 | 0.0700 | 2004-2017 | Just some interesting values |  |
| < 16 | 526 | 0.1300 | 2004-2017 | I understand that the minimum age to drive in Ontario is 16, that’s right? my idea to handle this is just removed the observations related |  |
| **ENVIRONMENT\_CONDITION (MVCR 0303)** | 00 - Unknown | 1541 | 0.4000 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **LIGHT (MVCR 0304)** | 00 - Unknown | 2692 | 0.7000 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **TRAFIC\_CONTROL (MVCR 0305)** | “” (Blank) | 49 | 0.0100 | 2014-2017 | Removed the observations related |  |
| 11 - Roundabout | 1007 | 0.2600 | 2014-2017 | There is a new version of the Manual? My idea to handle this is keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **TRAFFIC\_CONTROL\_CONDITION (MVCR 0306)** | "" (Blank) | 47022 | 12.3500 | 2014-2017 | Those observations could he homologated as "00 - Unknown"? |  |
| 00 - Unknown | 19250 | 5.0000 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| 88 | 131525 | 34.5000 | 2004-2013 | There is a new version of the Manual? My idea to handle this is keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **ROAD\_CHARACTER (MVCR 0307)** | This feature value will be ignored/removed based on information provided (no longer used) | | | | |  |
| **ROAD\_SURFACE (MVCR 0308)** | This feature value will be ignored/removed based on information provided (no longer used) | | | | |  |
| **ROAD\_CONDITION (MVCR 0309)** | 00 - Unknown | 5128 | 1.3000 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| 88 | 1 | 0.0003 | 2011 | Removed the observations related |  |
| **ROAD\_SURFACE\_CONDITION (MVCR 0310)** | 00 - Unknown | 5092 | 1.3000 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **ROAD\_ALIGNMENT (MVCR 0311)** | 00 - Unknown | 4829 | 1.3000 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **ROAD\_PAVEMENT\_MARKINGS (MVCR 0312)** | 00 - Unknown | 6464 | 1.6900 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **VEHICLE\_CONDITION (MVCR 0317)** | "" (Blank) | 10 | 0.0020 | 2015-2016 | Those observations could he homologated as "00 - Unknown"? |  |
| 13 - Other | 349 | 0.0900 | 2004-2017 | There is a new version of the Manual? My idea to handle this is keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| 99 - Defect | 96 | 0.0200 | 2015-2017 | There is a new version of the Manual? My idea to handle this is keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **DRIVER\_CONDITION (MVCR 0319)** | 08 - Inattentive | 26027 | 6.8000 | 2014-2017 | I already homologate it to "08 - Hand Held Device"(this supposed to be a NOT ALLOWED VALUE) |  |
| **INITIAL\_DIRECTION\_OF\_TRAVEL (MVCR 0323)** | "" (Blank) | 3 | 0.0008 | 2015 | Those observations could he homologated as "00 - Unknown"? |  |
| 00 - Unknown | 7619 | 2.0000 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **INITIAL\_IMPACT\_TYPE (MVCR 0324)** | "" (Blank) | 1 | 0.0003 | 2015 | Removed the observations related |  |
| **VEHICLE\_MANEUVER (MVCR 0325)** | "" (Blank) | 13 | 0.0030 | 2015 | Those observations could he homologated as "00 - Unknown"? |  |
| **VEHICLE\_DAMAGE (MVCR 0331)** | "" (Blank) | 62 | 0.0163 | 2014 - 2015 | Those observations could he homologated as "00 - Unknown"? |  |
| 00 - Unknown | 27831 | 7.3143 | 2004-2017 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **SEQUENCE1 (MVCR 0326/0327/0328/0329)** | I understand that this value could not be BLANK (NULL), that's right? | | | | |  |
| "" (Blank) | 1 | 0.003 | 2015 | Those observations could he homologated as "00 - Unknown"? |  |
| 00 - Unknown | 21 | 0.0055 | 2004-2013 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| 67 – Fire Hydrant | 91 | 0.0239 | 2004-2017 | There is a new version of the Manual? My idea to handle this is keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **SEQUENCE2 (MVCR 0326/0327/0328/0329)** | I understand that this value could be BLANK (NULL), that's right? | | | | |  |
| What is the difference between "00 - Unknown", "" (Blank) and "-1"? I guess that "-1" and "" (Blank) is something like "Not Apply", that's right? so in this case I can  homologate "" (Blank) and "-1" to "-1 - Not Apply" or something like that? | | | | |  |
| "" (Blank) | 94864 | 24.9315 | 2014 - 2017 | Merge those two values with something like “-1 Not Apply” (this supposed to be a NOT ALLOWED VALUE) |  |
| -1 | 234856 | 61.7232 | 2004 - 2013 |  |
| 00 - Unknown | 5 | 0.0013 | 2007,  2010, 2011 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| 67 – Fire Hydrant | 121 | 0.0318 | 2004-2017 | There is a new version of the Manual? My idea to handle this is keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **SEQUENCE3 (MVCR 0326/0327/0328/0329)** | I understand that this value could be BLANK (NULL), that's right? | | | | |  |
| What is the difference between "00 - Unknown", "" (Blank) and "-1"? I guess that "-1" and "" (Blank) is something like "Not Apply", that's right? so in this case I can  homologate "" (Blank) and "-1" to "-1 - Not Apply" or something like that? | | | | |  |
| "" (Blank) | 104331 | 27.4195 | 2014 - 2017 | Merge those two values with something like “-1 Not Apply” (this supposed to be a NOT ALLOWED VALUE) |  |
| -1 | 262355 | 68.9502 | 2004 - 2013 |  |  |
| 00 - Unknown | 2 | 0.0005 | 2005, 2007 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| 67 – Fire Hydrant | 74 | 0.0194 | 2004-2017 | There is a new version of the Manual? My idea to handle this is keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **OFFSET1 (MVCR 0330)** | I understand that this value could be BLANK (NULL) because it depends of the collision, that's right? | | | | |  |
| What is the difference between "00 - Unknown", "" (Blank) and "-1"? I guess that "-1" and "" (Blank) is something like "Not Apply", that's right? so in this case I can  homologate "" (Blank) and "-1" to "-1 - Not Apply" or something like that? | | | | |  |
| "" (Blank) | 105238 | 27.6579 | 2014 - 2017 | Merge those two values with something like “-1 Not Apply” (this supposed to be a NOT ALLOWED VALUE) |  |
| -1 | 258826 | 68.0228 | 2004 - 2013 |  |
| 00 - Unknown | 6579 | 1.7290 | 2004 - 2013 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **OFFSET2 (MVCR 0330)** | I understand that this value could be BLANK (NULL) because it depends of the collision, that's right? | | | | |  |
| What is the difference between "00 - Unknown", "" (Blank) and "-1"? I guess that "-1" and "" (Blank) is something like "Not Apply", that's right? so in this case I can  homologate "" (Blank) and "-1" to "-1 - Not Apply" or something like that? | | | | |  |
| "" (Blank) | 103026 | 27.0765 | 2014 - 2017 | Merge those two values with something like “-1 Not Apply” (this supposed to be a NOT ALLOWED VALUE) |  |
| -1 | 260133 | 68.3663 | 2004 - 2013 |  |
| 00 - Unknown | 3764 | 0.9892 | 2004 - 2013 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |
| **OFFSET3 (MVCR 0330)** | I understand that this value could be BLANK (NULL) because it depends of the collision, that's right? | | | | |  |
| What is the difference between "00 - Unknown", "" (Blank) and "-1"? I guess that "-1" and "" (Blank) is something like "Not Apply", that's right? so in this case I can  homologate "" (Blank) and "-1" to "-1 - Not Apply" or something like that? | | | | |  |
| "" (Blank) | 105653 | 27.7670 | 2014 - 2017 | Merge those two values with something like “-1 Not Apply” (this supposed to be a NOT ALLOWED VALUE) |  |
| -1 | 267373 | 70.2690 | 2004 - 2013 |  |
| 00 - Unknown | 1252 | 0.3290 | 2004 - 2013 | Keep it (this supposed to be a NOT ALLOWED VALUE) |  |