AUDIT REFERENCE: DRDGOLD SNOWFLAKE DATA INTEGRITY (YELLOW MACHINES USE CASE)

Queries Commenced 1st Dec 2024 and Terminated on 23Jan 2024 -

Dashboarding and Report Compilation commenced 03 Jan 2024 and terminated 28 Jan 2024

Data Sample Used (July – November 2024) ; Warehouse: TEST\_YMH, Database:CYBER\_SEC\_TESTING

5 Tables used: PorMasterDetail, QUARTZ/KLT/ REDTOP TELEMATICS, TIMESHEET\_TIMESHEET, TIMESHEET\_DAYS, Gm\_Details…

**Note: New vehicle tables were supplied at end of study, not incorporated in the report but not expected to change results drastically – it will need validation, but the information there in should have been incorporated in the tables used as the timestamp was April 2024, and this exercise commenced in November period of 24.**

Prepared By: GRAHAM BORTZ

Reviewed By:

Evidence Reference:

|  |  |
| --- | --- |
| FINDING DESCRIPTION  "PorMasterDetail\_PROD\_ERGO"  has ambiguous descriptions. | FINDING DETAIL  Many descriptions for the same vehicle leading to ambiguities and errors when trying to align this table with timesheets |
| CAUSE  Free text entry | DETAILED DESCRIPTION OF CAUSE.  The entry to the field describing the vehicle is via a text field (alphanumeric) and there does not appear to be a single description allowable for the vehicle. |
| RECOMMENDATION  Validation of entry | DETAILED REMEDIATION REQUIRED  There needs to be a front end to the entry mechanism. If this front end is internal to the system, then it needs to be a dropdown list where the data capturer can choose from predetermined choices. |
| MANAGEMENT COMMENT | MANAGEMENTS COMMENT RELATING TO FINDING |
| EVIDENCE TO SUPPORT FINDING  The evidence is given by a sample from the field in the database and calculations of the number of MStockCodes (Machine Descriptions) Per Machine | DETAILED DESCRIPTION OF EVIDENCE AND REFERENCE TO FINDINGS:  See the table below for the number of stock codes describing so few vehicles    Figure 1. The above shows there are 53 StockCodes describing 6 machine types., resulting in 62 Machine/Stock Code combinations.  This is broken down further by the actual machine below:  When the orders are joined to the time sheet for a single line order should describe one type of vehicle, but it does not seem to be the case. Below is evidence of a mismatch, possibly due to a description problem or an incorrect classification.    Figure 2. As can be shown above the Articulated Dump Truck is shown as a ‘Hire 18000L Water Cart’.  Dates are shown in the descriptions. This should not occur. In the same resulting table below, the vehicle type was left out, meaning that the timesheet never had a Machine filled in. This should not be possible.  Figure 3. It can be seen that in a few cases the machine was not provided as part of the timesheet data.  Now one can see evidence of many descriptions describing one machine. This is ambiguous as how does one know what machine the description alludes to. See the figure below.    Figure 4. Evidence of 1 description describing many Machines ie there are 14 stock codes for a front end loader - its unlikely that there are 14 different front end loaders. There are 14 descriptions for an ADT. This can lead to billing problems.  Below one can see one description describing multiple Machine Classes (see Hire 4x4 TLB – representing 3 classes.). One should also notice the slight differences in the spelling of the same description.    Figure 5. Each of the 4 x 4 TLB Codes describe multiple vehicle classes. Notice at the top the 3 different spellings of the same StockCode (Hire 4 x 4TLB, Hire 4 X 4 TLB, 4 X 4 TLB). Although unintentional these can describe different Machines as will be shown below:  The figure below shows that the ambiguous Stock Code descriptions describe different machines. This occurs multiple times. For example a Hire 4X4 describes tractors and front end loaders.    Figure 6. One stock code represents multiple machines, especially derivatives of 4X$ TLB. Even the 8 TON FEL , describer a tractor backhoe loader and a front end loader. |

|  |  |
| --- | --- |
| FINDING DESCRIPTION  The order table: "PorMasterDetail\_PROD\_ERGO"  has multiple line orders that have multiple vehicle in one order. | FINDING DETAIL  1) There is multiple line orders in the table mentioned: PorMasterDetail  2) More than 1 line in some of these orders contains a request for vehicle orders.  3) This creates a challenge to join the orders table to the timesheets table and track the timesheet hours against the cost awarded to the order.  3) It is also challenging to check whether the order amount conforms to the cost of the vehicle actually supplied.  4) It ‘seems that’ in order to resolve the ambiguity, descriptions for the same vehicle, as in the single line orders are being changed with added information (dates are added as well as vehicle codes). This again leads to inconsistent descriptions. |
| CAUSE  The DRD personnel constructing the orders can fill in multiple line orders. It seems that sometimes the orders are input in a way that may lead to ambiguities. | DETAILED DESCRIPTION OF CAUSE.  Multiple line orders contain orders for vehicles on more than 1 line. If there are rules then these are not being enforced.  The persons compiling the orders are using free text to complete the orders. |
| RECOMMENDATION  Define a rule for filling in multiple line orders | DETAILED REMEDIATION REQUIRED  Enforce a rule that multiple line orders can only order a vehicle hire on one of the lines. Other lines can contain requests that are unrelated to those that have timesheet requirements.  Do not allow free text entry, use validation or drop-down lists.  Multiple line orders for multiple vehicles per order can be accommodated in a different table going forward, with additional columns, Month, Date of Order, and Vehicle Code. This would allow the join to the timesheet data and resolve the ambiguity.  For clarity the term MStockCode should be a code in the strict sense of the word. The manner in which the codes are being inserted means that the code changes per month for the same vehicle and the code changes based on the vehicle code. The inconsistency (see evidence presented below) is that the same vehicle is described by month sometimes (no year) , and by Month and Year, or by vehicle code (no month and year). This is multiplying up the number of codes increasing the ambiguity |
| MANAGEMENT COMMENT | MANAGEMENTS COMMENT RELATING TO FINDING |

|  |  |
| --- | --- |
| EVIDENCE TO SUPPORT FINDING | DETAILED DESCRIPTION OF EVIDENCE AND REFERENCE TO FINDINGS  The table below shows evidence captured from a multiline order extraction:  Figure 7. Evidence of the aforementioned multiple line examples where multiple vehicles are ordered on different lines, there are 44 more rows of multiline orders.  As one can see from Figure 7 above the same vehicle is ordered with different descriptions based on the month it is ordered (see top 3 lines) or based on the Vehicle\_Code.  Note the above is a snapshot of the data but here is evidence that some of the above vehicle codes exist in the timesheets but not in the TELEMATICS DATA, or do not exist in the timesheets, ie SHI1 and HI20. A detailed set of findings on this topic is presented later.  The table below shows the orders with multiple lines, and these orders have multiple vehicles. Only the lines with vehicle hire have been included here.    Figure 8. Here is evidence of multiple vehicles hires on different lines of the same order.  As one can see above there are up to 6 vehicles ordered in the same order on different lines. There are 20 orders that have this occurrence, so it’s not just a single instance.  The effect of the escalation of stock codes; there are at this stage 44 stock codes describing 7 or 8 vehicle types. The stock codes that represent the same vehicle are very similar but not identical, due to free text insertions. Note how many times the stock codes are used in the first 10 or so lines below:    Figure 9. Number of times stock codes are used  If the stock codes are used once for a particular order, then it becomes a description and not a code. Codes should be used multiple times.  If the codes are used in a disciplined manner then if one orders the same vehicle it should not make a difference, description wise. Shown below are the similarities or differences between the stock codes in ordering vehicles in multiline orders vs single line orders:  There are 41 rows of instances where the description of a vehicle is not the same as that in the single line order. There are therefore 41 customizations of codes out of 44 codes in all that do not appear in the single lie orders.    Figure 9a. There are 41 cases of vehicle descriptions that appear  There are only 3 cases where they appear in both tables. See below.    Figure 9b. There are 3 cases of vehicle descriptions that are common between multiple line orders and single line orders |

|  |  |
| --- | --- |
| FINDING DESCRIPTION  Vehicle codes are missing between timesheets and supplier tables ie the telematics tables, as of the data supplied (static) | FINDING DETAIL  There are vehicle codes that are present in the timesheets that are not present in the telematics data (referred to as suppliers in these results) |
| CAUSE  The timesheet controller is not checking the validity of the code designation | DETILED DESCRIPTION OF CAUSE  There may be a different identifier. Initially informed by Craft that the vehicle code should be used as the identifier and this should be correlated with the Registration Numbers in the Telematics Data. This was verified later by Craft that the vehicle code was the correct identifier. Later, well into the exercise it was recommended to use the IMEI code by DRD as the identifier. However there is no IMEI code in the timesheet, and strictly speaking the vehicle code is used throughout these tables which could have repercussions, should the results of this exercise be accurate in the present tables (after this exercise).  The suggestion was to do an analysis of IMEI code verse vehicle mapping throughout time, but within the limitations of. An estimation of the results of such an exercise was computed with some interesting findings. This will be alluded to in the next section on this topic. |
| RECOMMENDATION  These vehicle code to registration number mappings should be consolidated | DETAILED REMEDIATION REQUIRED  The tables supplied by the suppliers should be validated against all the codes used in the orders and timesheets. For the interim, one suggestion could be that the IMEI number is joined to the vehicle code and that become a new identifier that is used throughout. It should be generated on the fly ie from two columns in the table, one of IMEI and the other of vehicle code from timesheets and the same in telematics. Should this result be changed on computation of this field, it would be detected and an alarm is triggered to suspend the order or acceptance of timesheet. |
| MANAGEMENT COMMENT | MANAGEMENTS COMMENT RELATING TO FINDING |
| EVIDENCE TO SUPPORT FINDING  Matching / Mismatching of vehicle codes and registration numbers when comparing those in timesheets with the supplier tables | DETAILED DESCRIPTION OF EVIDENCE AND REFERENCE TO FINDINGS  As far as the data supplied is concerned, there are some mismatches from each supplier’s table and the corresponding timesheets when comparing registration numbers and vehicle codes. What is concerning is that KLT in particular has 2 vehicle codes that are in the timesheets and not in the KLT telematics data. For KLT there are no vehicles in the timesheets that have a match to a corresponding vehicle in the supplier table. (telematics data).    Figure 10. Shows the matches and non matches of the vehicle codes in the timesheets to those in the relevant telematics data.  Quartz is the most aligned with 2 out of 21 missing supplier matches (10%).  Redtop has a significant number of vehicles in the timesheets that can’t be mapped to its telematics data (11/26 or 42%)  The table below identifies the supplier which has timesheets allocated to then that are not in the supplier table.    Figure 11. The table above shows the vehicle list attributed to each supplier which, as far as this dataset is concerned, has timesheets associated with the vehicle which can’t be found in the telematics data from that supplier |

|  |  |  |
| --- | --- | --- |
| FINDING DESCRIPTION  IMEI numbers not fixed to vehicle code / registration number | FINDING DETAIL  There is evidence, based on the data available at the time of the report, that some vehicles are associated with multiple IMEI numbers. | |
| CAUSE  The vehicle supplier is using multiple IMEI numbers with a particular vehicle | DETAILED DESCRIPTION OF CAUSE  There are 2 reasons for this:  The supplier is swapping it out. He may have a legitimate reason, or he is not conforming to the rules.  There is no automated detection mechanism | |
| RECOMMENDATION  There should be a detection system that would detect this behavior | DETAILED REMEDIATION REQUIRED  Either a sensor could be deployed and fixed to the vehicle that would detect the IMEI (RFID)  Data validation in the database could used (Stored procedure) to run through any table and ensure that the registered mapping of Vehicle ID to IMEI number is maintained.  Note: The fact that this is happening means that the IMEI number alone cannot be used to identify the vehicles in place of the Registration Numbers | |
| MANAGEMENT COMMENT | MANAGEMENTS COMMENT RELATING TO FINDING | |
| EVIDENCE TO SUPPORT FINDING  The queries processed show the prevalence of multiple IMEIs associated with a few vehicles. | DETAILED DESCRIPTION OF EVIDENCE AND REFERENCE TO FINDINGS  An exercise was done to capture the IMEIs from the suppliers table as well as the dates associated with the IMEI status.  The tables of this IMEI status capture are shown below for 2 suppliers where evidence of the behavior occurred.  Figure 12 IMEI status capture for Redtop vehicles    Figure 13 IMEI status capture for Quartz vehicles  The number of IMEIs per vehicle are now captured at the various status dates for RedTop vehicles.  Figure 14. The above figure shows that t a vehicle code that had multiple IMEI numbers.  Vehicle EE2 presented multiple IMEI numbers. RedTop had been informed to refrain from doing this. It was fortunate that the changes reflected during the period of this audit else this would not have been detected. Therefore it is possible that this may still be occurring more often. It may not be visible over this time period.  The next table shows the IMEI numbers for all those vehicles that had more than 1 IMEI number. The only vehicle that had more than 1 IMEI number was EE1 which had 3 IMEI numbers as shown in the table below.  Figure 15. The IMEI numbers for vehicle EE2 and the associated dates of status capture  There was a month (August 24) where EE1 exhibited 3 different IMEI numbers. This makes it very difficult to track this vehicle as this IMEI number may have been elsewhere (different company) and it may be the case that DRD was billed for time spent elsewhere. **Note that this behaviour may have occurred more regularly in the past or after the dataset was provisioned**.  It was also useful to see if the same IMEI number was associated with 2 different vehicles, which can also be a problem. As shown below this did not happen for Redtop **in the timespan of this dataset. One cannot say that this did not happen prior to July 2024 and that it did not happen post October (November being the last date of capture).**      Figure 16. The IMEI numbers for this period have only been present once, one of each vehicle. The result was ordered from Max count to Min count, so Max count was 1.  The exercise was repeated for Quartz where again 2 vehicles were found to exhibit this behaviour.  Figure 17. The query results show that 2 vehicles has 2 IMEIs during the period under investigation.  The IMEIs and the associated vehicles and months where the status was captured are shown below.    Figure 18. Multiple IMEIs Per Vehicle in the Quartz table  It can be seen above that Q13 changed IMEIs between July and August 2024 and Q27 changed IMEIs in November 2024 | |
| FINDING DESCRIPTION  There are timesheets that are in the Timesheet Days Table that do not reflect in the Timesheet Timesheet table | | FINDING DETAIL  A study was done on the mapping of Timesheet\_Days (timesheet\_ID) to Timesheet\_Timesheet (ID).  There should be a mapping (1:1), however at the time of analysis there were a few timesheets that were initiated prior to registering the timesheet in the Timesheet\_Timesheet worksheet. |
| CAUSE  A timesheet could have been compiled at the start of the contract, but was not registered in the Time\_sheet table at the time of data analysis | | DETAILED DESCRIPTION OF CAUSE  The details of why the mismatch happened could either be a timing issue, a process deficiency or that this is normal at the start of the contract / order. |
| RECOMMENDATION  A system/process rule needs to be put in place to prevent such. | | DETAILED REMEDIATION REQUIRED  This rule needs to be established in the database that a timesheet days ID cannot be registered without it appearing in the Timesheet\_Timesheet table. A stored procedure can be used to check this daily. |
| MANAGEMENT COMMENT | | MANAGEMENTS COMMENT RELATING TO FINDING |
| EVIDENCE TO SUPPORT FINDING  There are timesheet numbers in the Timesheet Days table that are not in the Timesheet\_Timesheet table | | DETAILED DESCRIPTION OF EVIDENCE AND REFERENCE TO FINDINGS  The table below shows the Timesheet\_IDs in the Timesheet\_Days table that are not in the list of IDs in the Timesheet\_Timesheet table.    Figure 19. The Timesheet IDs that are not in the Timesheet\_Timesheet table but are in the Timesheet\_Days table.  It is not clear if a rule has been set for this but it can lead to billing for hours that have not been accumulated in the Timesheet\_Timesheet table.  It may also be the case that because the data is a snapshot in time, the reconciliation may still occur, |

|  |  |
| --- | --- |
| FINDING DESCRIPTION  Vehicle IDs in the Timesheets do not always reconcile with registration numbers | FINDING DETAIL  The previous study indicated that there were many instances of Vehicles that are in the timesheets that are not in the supplier table. Strictly speaking this is not ALWAYS the case as will be shown in the detailed study below. |
| CAUSE  The cause could be from either side or both sides ie from the supplier side and / or from DRD side. | DETAILED DESCRIPTION OF CAUSE  Both DRD and the Suppliers (Redtop/Quartz/KLT ) need to reconcile the Vehicle\_Codes in the Timesheet\_Timesheet table with the Registration\_Number in the Suppliers table. |
| RECOMMENDATION  Reconciling the Registration number with the Vehicle Code | DETAILED REMEDIATION REQUIRED  Both DRD and the Suppliers (Redtop/Quartz/KLT ) need to reconcile the Vehicle\_Codes in the Timesheet\_Timesheet table with the Registration\_Number in the Suppliers table ie \*\*\*\_Telematics. An agreement needs to be made on a standardized registration number and a lookup table needs to be materialized in both databases where the incorrect codes are replaced by the agreed upon codes.  Please note: the descriptions have to be exact to be processed by normal queries and not include a sophisticated process (NLP) to obfuscate the errors in the data. Should this be done then it should be transparent and remedied in the tables in another column and not used exclusively in a backend process. |
| MANAGEMENT COMMENT | MANAGEMENTS COMMENT RELATING TO FINDING |
| EVIDENCE TO SUPPORT FINDING  The following tables will provide the findings | DETAILED DESCRIPTION OF EVIDENCE AND REFERENCE TO FINDINGS  The following table shows the relevant Redtop Registration numbers that were deemed to be in Redtop but not in the Timesheets. There may be 1 or 2 errors in the analysis based on the new data which hasn’t been incorporated but there is sufficient evidence here that there is a systemic problem. The discussion about the findings is on the next page    Figure 20 a. Vehicle Registration numbers and 20b. Vehicle codes on the right  Discussion on the above 2 tables:  The missing vehicles in the timesheets table (on the right) will be discussed with respect to their presence or not in the Redtop Telematics data on the left.  **Missing Vehicle Codes from Table on the Right: Timesheet**  BT10: Missing in the Redtop Table on the Left  C15: Missing in the Redtop on the Left  EL2 : Unlikely to be missing (see EL 2 on the left table)  PC350: Unlikely to be missing (See PC 350 on the left table)  R16: Unlikely to be missing (see R16 Bell ADT on the left table)  R17: Unlikely to be missing (see R17 / R17 BELL ADT ambiguity, left table)  R23: Unlikely to be missing (see R23 ADT Water Cart, left table)  R24: Unlikely to be missing (see R 24 Bell, left table)  R31 and R55 are missing in the Redtop Table |