

Reported Diversion Statistics

Data Processing Module:

ExpectedDemand_ExceedsFV_UnitConversion_UseVsDiv_Statistics.xlsx

Description: Calculate various statistics from reported diversion and use to gain insight on expected diversion amounts, and identify outlier reports that may be in error

Example: QAQC staff quickly identify a water right with a relatively large average monthly standard deviation. Further review shows there was 1 report with substantially different values than the rest of the reports, and it seems to be reported in the wrong units. Staff correct the errant records and move on to the next record.

What (is being flagged): Average monthly reported values of direct diversion, and diversion to storage, monthly standard deviations (by month, across reporting years), and average monthly standard deviation.

Why: Pre-calculate several classic statistical metrics to assist staff in identifying reporting outliers, to prioritize and target further investigation as to whether a report is in error or not, and make changes as needed.

How: Reshape the RMS reported data into rows of reporting years with columns of months, for direct diversion and diversion to storage. Then calculate monthly standard deviations (by column for each water right) to provide the value of a standard deviation across various reports for a given month. Then calculate the average of all months for each water right. While the average standard deviation is not a standard statistical metric, it is useful in collapsing all the statistical calculations into 1 value for each water right, which can be sorted to assist prioritizing review of water right reports.

Resolution: If it is determined the outlier is due to a reporting error, the errant reports can be overridden by QAQC staff to better represent actual/expected diversion.

Data Source(s):

ewrims_flat_file.csv

water_use_report.csv

Existing Fields:

APPL_ID - Data from water_use_report.csv

YEAR - Data from water_use_report.csv

MONTH - Data from water_use_report.csv

AMOUNT- Data from water_use_report.csv

DIVERSION_TYPE - Data from water_use_report.csv

APPLICATION_NUMBER – Data from ewrims_flat_file.csv

INI_REPORTED_DIV_AMOUNT – Data from ewrims_flat_file.csv

INI_REPORTED_DIV_UNIT – Data from ewrims_flat_file.csv

FACE_VALUE_AMOUNT – Data from ewrims_flat_file.csv

FACE_VALUE_UNITS – Data from ewrims_flat_file.csv

New Fields:

REPORTING_YEAR

INDEX

APPLID

JAN_DIRECT_DIVERSION

FEB_DIRECT_DIVERSION

MAR_DIRECT_DIVERSION

APR_DIRECT_DIVERSION

MAY_DIRECT_DIVERSION

JUN_DIRECT_DIVERSION

JUL_DIRECT_DIVERSION

AUG_DIRECT_DIVERSION

SEP_DIRECT_DIVERSION

OCT_DIRECT_DIVERSION

NOV_DIRECT_DIVERSION

DEC_DIRECT_DIVERSION

JAN_STORAGE_DIVERSION

FEB_STORAGE_DIVERSION

MAR_STORAGE_DIVERSION

APR_STORAGE_DIVERSION

MAY_STORAGE_DIVERSION

JUN_STORAGE_DIVERSION

JUL_STORAGE_DIVERSION

AUG_STORAGE_DIVERSION

SEP_STORAGE_DIVERSION

OCT_STORAGE_DIVERSION

NOV_STORAGE_DIVERSION

DEC_STORAGE_DIVERSION

ANNUAL_DIRECT

ANNUAL_STORAGE

DUPLICATE_DIRECT_STORAGE

CALENDAR_YEAR_TOTAL

MAY_TO_SEP_TOTAL_DIVERSION

FaceValue

Initial_Reported_Diversion
Units_IniDiv
IniDiv_Converted_to_AF
AnnualTotalDiversion
Annual_Use
Diversion_as_Percent_of_FV
Amount_over_FV
Diversion_as_Percent_of_IniDiv
Amount_over_IniDiv
Annual_Diversion_if_reported_in_Gallons
Annual_Diversion_if_reported_in_GPM
Annual_Diversion_if_reported_in_GPD
Annual_Diversion_if_reported_in_CFS
UNIQUE_APPL_ID
JAN_AVERAGE_DIRECT_DIVERSION
FEB_AVERAGE_DIRECT_DIVERSION
MAR_AVERAGE_DIRECT_DIVERSION
APR_AVERAGE_DIRECT_DIVERSION
MAY_AVERAGE_DIRECT_DIVERSION
JUN_AVERAGE_DIRECT_DIVERSION
JUL_AVERAGE_DIRECT_DIVERSION
AUG_AVERAGE_DIRECT_DIVERSION
SEP_AVERAGE_DIRECT_DIVERSION
OCT_AVERAGE_DIRECT_DIVERSION
NOV_AVERAGE_DIRECT_DIVERSION
DEC_AVERAGE_DIRECT_DIVERSION
JAN_AVERAGE_STORAGE_DIVERSION
FEB_AVERAGE_STORAGE_DIVERSION
MAR_AVERAGE_STORAGE_DIVERSION
APR_AVERAGE_STORAGE_DIVERSION
MAY_AVERAGE_STORAGE_DIVERSION
JUN_AVERAGE_STORAGE_DIVERSION
JUL_AVERAGE_STORAGE_DIVERSION
AUG_AVERAGE_STORAGE_DIVERSION
SEP_AVERAGE_STORAGE_DIVERSION
OCT_AVERAGE_STORAGE_DIVERSION
NOV_AVERAGE_STORAGE_DIVERSION
DEC_AVERAGE_STORAGE_DIVERSION
JAN_EXPECTED_TOTAL_DIVERSION
FEB_EXPECTED_TOTAL_DIVERSION
MAR_EXPECTED_TOTAL_DIVERSION
APR_EXPECTED_TOTAL_DIVERSION
MAY_EXPECTED_TOTAL_DIVERSION

JUN_EXPECTED_TOTAL_DIVERSION
JUL_EXPECTED_TOTAL_DIVERSION
AUG_EXPECTED_TOTAL_DIVERSION
SEP_EXPECTED_TOTAL_DIVERSION
OCT_EXPECTED_TOTAL_DIVERSION
NOV_EXPECTED_TOTAL_DIVERSION
DEC_EXPECTED_TOTAL_DIVERSION
ANNUAL_TOTAL_DIVERSION
MAY_TO_SEP_TOTAL_DIVERSION
TOTAL_ANNUAL_USE
JAN_STDEV
FEB_STDEV
MAR_STDEV
APR_STDEV
MAY_STDEV
JUN_STDEV
JUL_STDEV
AUG_STDEV
SEP_STDEV
OCT_STDEV
NOV_STDEV
DEC_STDEV
AVERAGE_STDEV
Total_Cumulative_Diverted
Total_Cumulative_Use
Total_Use_as_a_Percent_of_Total_Diverted