



IT Collect "How To" Series

How to use the Data Quality Report

October 2022

v2.0

General Services Administration
Office of Government-wide Policy

Document Change Record

Version Number	Date	Description
1.0	September 2021	Created IT How-To Data Quality Report document outlining IT Collect data reports.
1.1	October 2021	Added additional summary information for Data Quality Report.
2.0	October 2022	Updated with new BY 2024 Statistics Reports.

What is the IT Collect Statistics service?

The IT Collect Statistics service is a GET-only, multi-report mechanism that displays data quality, data completeness, data staleness, and other metrics within reports to agencies. This series will focus on the data quality report that allows users to better understand the composition of their investment data. Some portions of the data quality report contain their own sub-reports (Result Sets) that provide additional data visibility.

What is displayed in the Data Quality Report?

Stock Data Quality Report Endpoint

This collection reports on the quality/completeness of the data submitted to the IT Collect database for a given agency. These reports are designed to be viewed by agency end users via a data feed on the modernized IT Dashboard front-end.

Agencies and relevant ITPfM stakeholders can leverage the Data Quality Report to help ensure the completeness and validity of submitted ITPfM data to the IT Dashboard. The Data Quality Report provides a variety of metrics to help Agencies verify that they have submitted all relevant data points. For example, one such Data Quality Report check will highlight all investments that have not provided projects data. The Data Quality Report also serves to highlight data instances that potentially fall outside expected norms of ITPfM data. An example of such a check would include highlighting new investments that contain PY or CY funding.

The Data Quality Report is designed to assist Agencies monitor their submitted data, and any Data Quality Report finding does not necessitate action on its own. Agencies are not obliged to take direct action solely to "resolve" Data Quality report findings. In other words, just because a data element is flagged in the Data Quality Report does not mean an Agency must or needs to "fix" the underlying data.

All reports apply to unretired UII's (isRetired = false) and non-Budget Authority Ledger Element Styles (ledgerElementStyle != "Budget Authority"). Below are the available filters, individual result sets, their subsets, and descriptions agencies will see in this report:

Filter	Description	Inputs
report	Required filter which allows users to designate which report to run	All available Result Sets outlined below
page	Specifies which page number a user wants to display. The first page will display 200 records, the second the next 200, and so on.	Integer value

BY 2024 Adjustments: Rows highlighted in green have been added for BY 2024. Whereas rows highlighted in red have been removed for BY 2024

Result Set	Subsets	Description	Fields Returned
missingCioRatings	N/A	<i>UII's that do not have CIO Ratings and have investmentType = 01 or 05</i>	agencyId (service) agency currentUII
missingContracts	N/A	<i>UII's that do not have Contracts and have investmentType = 01, 02, or 05</i>	agencyId (service) currentUII
missingProjects	N/A	<i>UII's that do not have Projects and have investmentType = 01</i>	agencyId (service) currentUII
activityStartDateTooEarly	N/A	<i>Activities with any start date before their parent project start date</i>	agencyId (activity) plannedStartDate projectedStartDate actualStartDate agencyId (service) currentUII AgencyId (project) StartDate (project)
activityEndDateTooLate	N/A	<i>Activities with any completion date following their parent project completion date</i>	agencyId (activity) plannedEndDate projectedEndDate actualEndDate agencyId (service) currentUII activityId (project) projectEndDate
newServicesWithCYPYFunding	N/A	<i>UII's with CY or PY funding AND changeInStatus = 10 (New Investment)</i>	year (ledger) amount (ledger) agencyId (ledger) currentUII agency agencyId (service)
eliminatedInvestmentsWithCYBYFunding	N/A	<i>UII's with CY or BY funding AND changeInStatus = 06, 07, 08, or 09 (Eliminated Investments)</i>	year (ledger) amount (ledger) agencyId (ledger) currentUII agency agencyId (service)
projectMissingActivities	N/A	<i>Projects with no activities</i>	agencyId (project) projectName agencyId (service) currentUII (service)
missingCostPools	missingAllCostPools	<i>UII's with no cost pools at all</i>	changeInStatus agencyId (service) currentUII

Result Set	Subsets	Description	Fields Returned
	missingPreviousYearCostPools	Existing UII's (changeInStatus != 10) missing PY cost pools	agencyId (service) changeInStatus currentUII
	missingCurrentYearCostPools	UII's that are not new or eliminated (changeInStatus != 10, 06, 07, 08, or 09) and are missing CY cost pools	agencyId (service) changeInStatus currentUII
	missingBudgetYearCostPools	UII's that are not eliminated (changeInStatus != 06, 07, 08, or 09) and are missing BY cost pools	agencyId (service) changeInStatus currentUII
missingITTowers	missingAllITTowers	UII's with no IT towers at all	changeInStatus agencyId (service) currentUII
	missingPreviousYearITTowers	Existing UII's (changeInStatus != 10) missing PY IT towers	agencyId (service) currentUII changeInStatus
	missingCurrentYearITTowers	UII's that are not new or eliminated (changeInStatus != 10, 06, 07, 08, or 09) and are missing CY IT towers	agencyId (service) currentUII changeInStatus
	missingBudgetYearITTowers	UII's that are not eliminated (changeInStatus != 06, 07, 08, or 09) and are missing BY IT towers	agencyId (service) currentUII changeInStatus
missingFundingSources	missingAllFundingSources	UII's with no funding (henceforth derived from ledger records containing funding source data) at all	changeInStatus agencyId currentUII
	missingPreviousYearFundingSources	Existing UII's (changeInStatus != 10) missing PY funding	agencyId (service) currentUII changeInStatus
	missingCurrentYearFundingSources	UII's that are not new or eliminated (changeInStatus != 10, 06, 07, 08, or 09)	agencyId (service) currentUII changeInStatus

Result Set	Subsets	Description	Fields Returned
		<i>and are missing CY Funding Sources</i>	
	missingBudgetYear FundingSources	<i>UII's that are not eliminated (changeInStatus != 06, 07, 08, or 09) and are missing BY Funding Sources</i>	agencyId (service) currentUII changeInStatus
cioRatingsInRedOverThreeMonths	N/A	<i>UII's with a red CIO rating (cioRating = 1 or 2) for more than 90 days</i>	title (service) currentUII agencyId (service)
cioRatingsInRedOverOneYear	N/A	<i>UII's with a red CIO rating (cioRating = 1 or 2) for more than 365 days</i>	title (service) currentUII agencyId (service) date (CIO rating) rating (CIO rating) agencyId (CIO rating)
missingOperationalAnalyses	N/A	<i>UII's that are missing operational analysis data and have investmentType = 01</i>	agencyId (service) currentUII
projectLifecycleLedgerExcess	N/A	<i>UII's where project lifecycle costs (projected total cost) exceed ledger lifecycle costs</i>	agencyId (service) currentUII title (service) projectLifecycleCost ledgerLifecycleCost
UIIsToBeRetired	N/A	<i>UII's that are eliminated (changeInStatus = 06, 07, 08, 09) and are not yet Retired (isRetired = false)</i>	changeInStatus currentUII title (service) agencyId (service)
activitiesOverOneYear	N/A	<i>Activities with a duration (planned/projected/ actual end date – planned/projected/ actual start date) exceeding one year</i>	currentUII agencyId (service) title (service) agencyId (project) projectName agencyId (activity) activityName actualStartDate, plannedStartDate, ProjectStartDate actualEndDate, plannedEndDate, projectedEndDate

Result Set	Subsets	Description	Fields Returned
projectsMissingLowestLevelActivities	N/A	<i>UII's that do not contain a single activity where lowestLevel = true</i>	agencyId (service) title (service) currentUII agencyId (project) projectId
lowestLevelValidityCheck	N/A	<i>(1) UII's where an activity has lowestLevel = true but has a sibling activity that has a lower structureId OR (2) activity has lowestLevel = false but does not have a sibling activity that has a lower structureId -- *Report can only be run on a single investment*</i>	agencyId (service) title currentUII agencyId (project) projectId agencyId activityName structureId providedLowestLevel expectedLowestLevel
majorInvestmentsMissingMetrics	N/A	<i>UIIs that do not have metrics data and have investmentType = 01</i>	agencyId (service) title (service) currentUII (service)
missingStrategicResultsMetrics	N/A	<i>UIIs with investmentType = 01 that do not have at least three isRetired = false metrics with performanceMeasurementCategory = "Strategic and Business Results"</i>	agencyId (service) title (service) currentUII (service)
missingStrategicResultsMonthlyMetrics	N/A	<i>UIIs with investmentType = 01 that do not have at least one isRetired = false metric with performanceMeasurementCategory = "Strategic and Business Results" and reportingFrequency = "Monthly"</i>	agencyId (service) title (service) currentUII (service)
missingCustomerSatisfactionMetrics	N/A	<i>UIIs with investmentType = 01 that do not have at</i>	agencyId (service) title (service) currentUII (service)

Result Set	Subsets	Description	Fields Returned
		<i>least one isRetired = false metric with performanceMeasurementCategory = "Customer Satisfaction"</i>	

❖ Here are two sample Result Sets users will see within the Data Quality Report. On the left is missingCioRatings and on the right is missingOperationalAnalyses. Users can expect all Result Sets under the Data Quality Report to be displayed similarly to these two examples:

```
{
  "@context": "/contexts/DataQualityReport",
  "@id": "/v1/data-quality-report",
  "@type": "hydra:Collection",
  "hydra:member": [
    {
      "_id": "01F3DCAH5RB6K20DXZFTHRDSV",
      "agency": "018",
      "agencyId": "string",
      "currentUII": "018-000000283"
    },
    {
      "_id": "01F3DCGJR3XJ70XH2JFR4G6N63",
      "agency": "018",
      "agencyId": "string",
      "currentUII": "018-000000293"
    },
    {
      "_id": "01F3DNVCECAY700V68S9WSHT3",
      "agency": "018",
      "agencyId": "string",
      "currentUII": "018-000000124"
    }
  ],
  "hydra:totalItems": 3,
  "hydra:view": {
    "@id": "/v1/data-quality-report?report=missingCioRatings",
    "@type": "hydra:PartialCollectionView"
  },
  "hydra:search": {
    "@type": "hydra:IriTemplate",
    "hydra:template": "/v1/data-quality-report{?report}",
    "hydra:variableRepresentation": "BasicRepresentation",
    "hydra:mapping": [
      {
        "@type": "IriTemplateMapping",
        "variable": "report",
        "property": null,
        "required": true
      }
    ]
  }
}
```

```
{
  "@context": "/contexts/DataQualityReport",
  "@id": "/v1/data-quality-report",
  "@type": "hydra:Collection",
  "hydra:member": [
    {
      "agencyId": "01F3BES0NQHJRHNSXNY0Q3HZR7",
      "currentUII": "018-000000003"
    },
    {
      "agencyId": "01F3BES0NWINTS3WCRNQWSYHXG",
      "currentUII": "018-000000004"
    },
    {
      "agencyId": "01F3BES0P2FQWBM31W91YB1J88",
      "currentUII": "018-000000005"
    }
  ],
  "hydra:totalItems": 3,
  "hydra:view": {
    "@id": "/v1/data-quality-report?report=missingOperationalAnalyses",
    "@type": "hydra:PartialCollectionView"
  },
  "hydra:search": {
    "@type": "hydra:IriTemplate",
    "hydra:template": "/v1/data-quality-report{?report}",
    "hydra:variableRepresentation": "BasicRepresentation",
    "hydra:mapping": [
      {
        "@type": "IriTemplateMapping",
        "variable": "report",
        "property": null,
        "required": true
      }
    ]
  }
}
```

Active Risk Report (BY 2024 Addition)

This report dynamically finds UIIs with investmentType = 01 that cross a user-defined threshold of active risk area totals. Users can select whether to this report will return UIIs that are either tracking a number of Risk Areas above or below the user provided threshold. Agencies can leverage this report to locate UIIs that might have erroneously omitted Risk Areas. Or, Agencies could locate UIIs tracking an above average number of Risk Areas for a give UII. Below are the available filters, individual result sets, and descriptions agencies will see in the report:

Filter	Description	Inputs
page	Specifies which page number a user wants to display. The first page will display 200 records, the second the next 200, and so on.	Integer value
activeRiskCountLess	Returns UIIs with a count of active Risk Areas less than the user provided value	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13
activeRiskCountMore	Returns UIIs with a count of active Risk Areas more than the user provided value	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 12

Result Set	Description
currentUII	<i>currentUII of UII</i>
title	<i>Title of UII</i>
agencyId	<i>Agency provided unique identifier of UII</i>
activeRiskCount	<i>Total count of active risk areas for a given UII</i>

TBM Discrepancy Report

This report displays a list of investments with either a cost pool or IT tower total for a given year that is greater than a plus or minus 5% delta than the funding amount total. All UII's displayed in this report are not retired (isRetired != false) and the funding amount totals do not include budget authority ledgers (ledgerElementStyle != "Budget Authority"). Below are the available filters, individual result sets, and descriptions agencies will see in the report:

Filter	Description	Inputs
page	Specifies which page number a user wants to display. The first page will display 200 records, the second the next 200, and so on.	Integer value

Result Set	Description
funding_total	<i>Total Funding amount (\$M) for a given year</i>
cost_pool_total	<i>Total Cost Pool amount (\$M) for a given year</i>
cost_pool_deviation	<i>% Difference between Cost Pool total and Funding total for a given year</i>
it_tower_total	<i>Total IT Tower amount (\$M) for a given year</i>
it_tower_deviation	<i>% Difference between IT Tower total and Funding total for a given year</i>

- ❖ Below is a sample report users will see for the TBM Discrepancy Report. Users can expect their report to be displayed in a similar way:

```

{
  "@context": "/contexts/TBMDiscrepancyReport",
  "@id": "/v1/tbm-discrepancy-report",
  "@type": "hydra:Collection",
  "hydra:member": [
    {
      "018-0000000000": {
        "2019": {
          "funding_total": 12.907,
          "cost_pool_total": 27.967,
          "cost_pool_deviation": 1.166809,
          "it_tower_total": 0,
          "it_tower_deviation": -1
        },
        "2020": {
          "funding_total": 5.57,
          "cost_pool_total": 0.6970000000000001,
          "cost_pool_deviation": -0.874865,
          "it_tower_total": 0,
          "it_tower_deviation": -1
        },
        "2021": {
          "funding_total": 50.14,
          "cost_pool_total": 17.869,
          "cost_pool_deviation": -0.643618,
          "it_tower_total": 58.183,
          "it_tower_deviation": 0.160411
        }
      }
    }
  ],
  "hydra:totalItems": 1
}

```

Portfolio Report

This report displays all non-retired (isRetired = false) investment-level UII's for each time period. This report is intended for agencies to see how many Services they have submitted across their portfolio. Here is an example of this report and available filters:

Filter	Description	Inputs
report	Only one result set is available for the Portfolio Report	PortfolioReport
page	Specifies which page number a user wants to display. The first page will display 200 records, the second the next 200, and so on.	Integer value

```
{
  "portfolioReport": {
    "TimePeriod20211": {
      "018": [
        "018-0000000000",
        "018-0000000001",
        "018-0000000002",
        "018-0000000003",
        "018-0000000004",
        "018-0000000005"
      ]
    },
    "TimePeriod20212": {
      "018": [
        "018-0000000000",
        "018-0000000001",
        "018-0000000002",
        "018-0000000003",
        "018-0000000004",
        "018-0000000005"
      ]
    }
  }
}
```