

# FEM Data Guide

November 04, 2021

## Contents

<b>1</b>	<b>Introduction to the Higg Facility Environmental Module</b>	<b>2</b>
1.1	About This Document . . . . .	2
<b>2</b>	<b>Getting Started With Higg FEM Data</b>	<b>2</b>
<b>3</b>	<b>Making Sense of Variable Names (REF IDs)</b>	<b>3</b>
3.1	Filters or Characteristics . . . . .	3
3.2	Scores . . . . .	3
3.3	Quantitative Data . . . . .	4
3.4	Self-Reported and Verified Data . . . . .	5
3.5	Section Questions . . . . .	5
3.6	Other suffixes and terms . . . . .	6
<b>4</b>	<b>Tips and Tricks for Analytics</b>	<b>6</b>
4.1	Year Over Year Comparison . . . . .	6
4.2	Comparing groups of facilities . . . . .	6
4.3	Further Analytics Support . . . . .	7
<b>5</b>	<b>Introduction to Documentation</b>	<b>8</b>
5.1	Applicabilities . . . . .	8
5.2	Questions (by Level) . . . . .	8
5.3	Tables . . . . .	8
5.4	Option Lists . . . . .	8

# 1 Introduction to the Higg Facility Environmental Module

The Higg Facility Environmental Module (FEM) measures annual facility environmental performance across seven areas:

- Environmental Management System (EMS)
- Energy/Greenhouse Gas Emissions
- Water
- Waste
- Wastewater
- Air Emissions
- Chemicals Management

## 1.1 About This Document

This is a reference guide for navigating Higg FEM data. It provides a technical, slightly narrative description of the Higg FEM data as well as reference documentation for all FEM questions (including tooltips, conditionalities, answer options, and points).

Some users may find it most useful to search keywords or variable names (called REF IDs) to find the corresponding question, answer options, or available points. Other users may find the section "Making Sense of REF IDs" helpful by breaking down how those variable names are constructed. And finally, some users will find the "Intro to Analytics" section useful for jump-starting Higg FEM analysis.

This document does not detail Higg FEM scoring procedures (it does show the points per scored question), nor offer insight into the intent of Higg FEM content. For the most publicly available information on scoring, see [this How-to-Higg page](#), while the latter can be found on [HowtoHigg.org](#). If you need help filling out a Higg FEM, [How-to-Higg](#) is your go-to resource, offering both question and content guidance. Users who would like the full list of variables in the bulk csv can download an excel guidance document [at the bottom of this webpage](#). For more advanced analytics support, please see the Higg package options on the [Higg Portal](#).

## 2 Getting Started With Higg FEM Data

The Higg FEM contains a ton of data. We know. If you download the full dataset, you'll find around 10k variables! And the Higg FEM is updated yearly, so it can be a challenge to stay current. With that in mind, how can you orient yourself to start performing analytics and making decisions with this veritable treasure trove of data? We're here to help.

First things first, where's the data? Users with Advanced Services can find adjustable dashboards to perform analyses on the Higg platform. Whether you upgrade or not, the data from facilities that have shared with you (or your data if you are a facility) [can be downloaded in csv form from the Higg Portal](#).

OK, we've got our data, what should we do? Well, you're in a great place. This document is your friend. You may also like to see the excel reference guide, which contains a list of every variable found in the bulk csv and can be downloaded at [the bottom of this webpage](#).

The Bulk CSV includes 6 key sections of information:

1. Account and module information

2. Higg FEM scoring results for total, section, level, and question scores
3. Air emission, wastewater, water and chemical management sections applicability
4. All module response data (answers to module questions)
5. All module verification data (including the verified result and corrected response)
6. Total GHG emission amount and GHG calculations

## 3 Making Sense of Variable Names (REF IDs)

Before we dive into that content documentation we're first going to talk about variable names. The key to any analysis is selecting the correct variables, so this will hopefully unlock a lot of possibility on its own.

A REF ID is shorthand for "Reference (REF) Identifier (ID)" and is a unique variable identifier. In addition to being a key part of the dataset, REF IDs are also found on the Higg FEM platform next to the question. Here is a guide for breaking down the naming conventions so you can quickly find the needed variables to make your analyses come to life. Once you're through this section, use control-F on this document to locate variables and their corresponding information (question, supporting info or tooltip, answer options, and possible points).

It's important to note that the bulk CSV download contains REF IDs for multiple versions of FEM. Some REF IDs may only be applicable to certain years of FEM, and will not be populated in assessments from other years.

### 3.1 Filters or Characteristics

The Site Information section contains key variables such as country, facility type, processes, materials, permits, and more. These can be used to filter the dataset or classify facilities.

These REFIDs typically have a prefix of "sip", which stands for "site information and permits". Examples include sipfacilitytype and sipmaterialplastics.

The applicability REF IDs are {section}.applicability. For example, "chemicals.applicability". More on applicabilities can be found later in the document.

### 3.2 Scores

Scores are one of the key performance metrics in the Higg FEM. Scores are generally applied to combinations of qualitative metrics (for example, Yes/No, or Was this quantitative value supplied?). Quantitative performance (for example, GHGs/kg production) are for the most part unscored.

To find the scores in the Higg FEM data:

#### 1. Question-level Scores

- Begin with "{section}\_". These sections are energy, water, air, wastewater, waste, ems, chemicals. These prefixes are followed by the refid to which they correspond.
- Ex. water\_watsources, wastewater\_wwsepticwater, and air\_airbeyondpermit

## 2. Level, Section, and Total Scores

- Level scores follow the pattern "{section}.level{number}", where number could be 1, 2, or 3.
- Level achievement follows the pattern "{section}.levelsAchieved"
- Section scores follow the pattern "{section}.score".
- The total FEM score is "score.total".

At this point, the SAC does not make public which questions contribute in which ways to create scores. The points per scored question are available in the tables below. To clarify this a bit more, while only one question is scored, the scoring rule is often comprised of many sub-questions. For the most publicly available information on scoring, see [HowToHigg](#).

## 3.3 Quantitative Data

The next type of data you will likely want to use in your analyses would be quantitative data. Keep in mind that units can differ within many of the quantitative datapoints. Units are often reported separately and will usually contain the term "unit" (sometimes "units") in the REFID, though occasionally it will be "type". For totals of sources, there will be consistent units.

The most highly sought after data are **Greenhouse Gas Emissions (GHGs)**. These REFIDs end in one of two ways: either ".GHG: renewable" or ".GHG: nonrenewable" (these will be e.g. ".GHG.verified: renewable" if verified data), where the second part denotes whether the energy source is considered renewable or nonrenewable. The first part of the REFID is "{fuel source category}{fuel source details}". For example, a complete REFID is "natgaslng.GHG: nonrenewable" for Liquefied Natural Gas. For total GHGs, use the variable "totalGHGEmissions". Please note that this figure excludes GHGs from biomass, in accordance with GHG Protocol standards with respect to short-term GHGs. All GHG REFIDs are reported in units of kilograms CO<sub>2</sub>eq (kg CO<sub>2</sub>eq).

The other quantitative data you may find most useful are:

- **energy use:** For the overall total, see "ensourcetotal". For specific source totals, they are written as "ensource{sourcename}total". For example, "ensourcecoaltotal". The source totals are reported in megajoules (MJ). Facilities can self-report their energy use in many different units, so if using those REF IDs, be sure to check the units.
- **water use:** For the overall total, see "watsourcetotal" and "watsourcetotaltotal". The former is for FEM2020+ data whereas the latter contains data prior to FEM2020. It is likely that you can safely merge these two columns to use as a total water consumption value across any year. For specific source totals, they are written "watsource{sourcename}total". For example, "watsourcemunicipaltotal". Source totals are reported in liters (L).
- **waste production:** For the overall total, see "wstsourcetotal". For aggregates of individual sources, see "wstsourcenhmaterialtotal" and "wstsourcenhtotaldomtotal". For specific source totals, they are written as "wstsource{nh/h}{sourcename}total". Note the {nh/h} (i.e., nonhazardous/hazardous) distinction. That is important for coding as well as for proper analyses and easy to miss! Source totals are reported in kilograms (kg).
- **emissions to air:** For overall totals, see "totalRefrigerants". This is reported in kilograms of CO<sub>2</sub>eq (kg CO<sub>2</sub>eq). Facilities are able to report up to five different sources of refrigerant emissions, and these REF IDs are written as "airsourcerefrig{1-5}total", which are reported in kilograms (kg).

- **wastewater production:** For overall totals, see "wwtrackindamt", "wwtrackdomamt", and "wwtrackcombamt". These are reported in cubic meters (m3).

The FEM also collects baselines and targets for many if not all of these same sources. Key terms to look for in REFIDs are "baseline" and "target". The standard questions that ask for source values will just have "{section shorthand}source", whereas a target will be "wsttargetnhsourcemetals".

Other tips for working with quantitative data:

- **Quantitative data** will often end in (or contain) one of the following terms: total, quant, amt, qty.
- The units that correspond to the source will often replace {total} with units or unit.
- **Quantitative data** also asks for the "method/s" and frequency ("freq") of data collection, as well as often a question whether a source is tracked ("track").
- Higg FEM energy data collection was streamlined in 2020. For instance, Higg FEM no longer asks for amounts of types of coal. However, there are still REF IDs for those variables in the dataset in order to contain legacy data (2017-2019).

### 3.4 Self-Reported and Verified Data

When you do ANY of these analyses on scores or quantitative data, you have two choices for suffixes that correspond to whether the data is self-reported or verified. If a facility has verified data, "verifiedposted" will be TRUE (or 1).

There are multiple ways self/verified data are distinguished in the Higg FEM. The base REF ID for both self-reported and verified values will always be the same, but they will have different **suffixes**.

1. **Self-reported values** will show up in several ways across the dataset. The most common suffix is ".value", which appears for user-inputted values. For calculated and derived values (like scores), the ref id will end in ".facility" and lack a suffix.
2. **Verified values** will appear as ".corrected" (for user-inputted values), while ".verified" appears for calculated and derived values. Analyses of verified data should use the ".corrected" value if populated.
3. **Selection values** are metadata that refer to whether a verifier flagged the initial user input as correct/incorrect. They are often inconsistently applied by verifiers, and can be disregarded for most analyses.

You may want to [download the excel-based csv reference guide at the bottom of this webpage](#), which contains the suffixes for every variable along with their definitions.

### 3.5 Section Questions

These REFIDS follow a similar pattern: "{section shorthand}{question shorthand}". The question shorthands are unique and give a (at times clear, at other times obscure) clue to the content of the question, however the section shorthands are consistent. The section shorthands for each section are:

- Energy - en
- EMS - ems

- Chemicals - chem
- Water - wat
- Waste - wst
- Wastewater - ww
- Air - air

An example REFID is "**wst**baselinedisp".

### 3.6 Other suffixes and terms

You may find yourself wanting to filter out or for other types of variables. For this, keep in mind common suffixes and terms:

- Fairly self-explanatory REFID identifiers include: year, note/s, comment/s, describe (and "description" and "desc"), and file.
- "agenc" which refers to regulatory agency.

## 4 Tips and Tricks for Analytics

Now that you have a sense for the variables (or at least the variable names) in the Higg FEM, you may have sense of what analyses you can run. In any case, here are a few tips and tricks.

### 4.1 Year Over Year Comparison

Some of the most common analyses are year over year comparison across scores and tracking data (e.g., emissions and water consumption). When doing so, keep in mind these several tips:

- Pay attention to applicabilities when working with scores! Applicabilities determine both the questions a user sees as well as the value of each scored question. Because every facility can receive the same total score, analyses of total score will not be as affected as analyses of particular scores.
- Pay attention to corresponding units when working with quantitative data! Facilities are able to report in whichever units are appropriate.
- Be aware of outliers! You may want to exclude values that fall 3 or 4 standard deviations from the mean or log transform the variables to get a better view. This is especially true for self-reported data.
- If a facility has not reported in multiple years, you can use their reported baseline data to see what progress has been made.

### 4.2 Comparing groups of facilities

Everyone will have a different way of grouping like facilities. You may prefer to compare or filter facilities across several characteristics such as type of facility processes, tiers, or countries.

The common grouping variables are:

- sipfacilitytype
- version
- country
- sipfacilityprocesses{sew/printdye/materials/trim/chem/pkg}
- sipfacilityfulltimeemployees
- sipfacilityannualprodvolquant (pay attention to sipfacilityannualprodvolunits)
- Wastewater standards and permits, which can be found in the Wastewater Applicabilities section below.

Here is a basic framework for grouping facilities into value chain tiers:

	REF ID	Tier
sipfacilitytypesew		1
sipfacilitytypeprint		2
sipfacilitytypematerials		2
sipfacilitytypetrim		2
sipfacilitytypepackaging		2
sipfacilitytypechem		4

### 4.3 Further Analytics Support

For advanced analytics support, you can purchase a package upgrade on the [Higg platform](#).

For information on how to use this data to improve facility performance, you can reach out to the [SAC](#) or the [Apparel Impact Institute](#).

## 5 Introduction to Documentation

### 5.1 Applicabilities

Five of the seven sections have applicabilities (not EMS nor Energy). In the applicabilities table at the beginning of the section, you will find the different possible applicabilities as well as the questions that determine which applicability-path a user will follow.

The Energy and Waste section do not have applicabilities, so in the points column you will find AppEn001 and AppWst001 for all surveys.

### 5.2 Questions (by Level)

The section questions are grouped by level. The details for each question are:

1. **REF ID** (i.e., Reference ID) is the column name in the csv export. It is your key to link the csv to the FEM survey.
2. **Question Text** is what a user sees in the Higg FEM survey.
3. **Answer Type**
4. **Conditional** displays if asking the question is conditional upon another questions answer. So if you see "emsmgmt,Yes" that means that this question is only asked if the user responded "Yes" to "emsmgmt".
5. **Option list** displays what option list corresponds to this particular question. The option lists contain all of the choices available to the user.
6. **Points** displays the points the question is worth depending on the applicability. If points differ based on applicability, the column will read "[Applicability1]: x; [Applicability2]: y". Scored questions are in bold and shaded in gray. See the **Points** column to determine whether it is scored for any particular applicability.

### 5.3 Tables

There are many "tables" of questions within the FEM, appearing as tables on the Higg interface rather than standalone questions. The main table REF ID will appear in the standalone question list, and the "Answer Type" will be "table". For example, "emsmgmttable" corresponds to EMS Management Table. The table REF ID can be found in the footnote to the table.

### 5.4 Option Lists

These are the answer options for each of the dropdown or multiselect option lists in a given section.