

DATA DICTIONARY – IT Metrics

Erik Brynjolfsson, Wang Jin, and Sebastian Steffen

Table 1: Summary of Datasets

File Name	Entity Level	Entity Concept	Example Entity Values	Time Level
naics2_it_metrics	2-Digit NAICS Industries	Industry	54 (Professional, Scientific, and Technical Services)	Year
naics6_it_metrics	6-Digit NAICS Industries	Industry	541330 (Engineering Services)	Year
firms_it_metrics	All Firms	Firm	Microsoft Corporation	Year
gvkey_it_metrics	Public Firms	Firm	12141 (Microsoft Corporation)	Year
state_it_metrics	States	Geography	Massachusetts	Year
fips_it_metrics	6-Digit FIPS Counties	Geography	25025 (Suffolk County, MA)	Year

VARIABLE DESCRIPTIONS

1. Occupation-based Variables

- **bgt_soc6_total**

The total number of job postings posted by an entity.

- **bgt_it_soc6_cyber_total**

The total number of cybersecurity-related job postings, based on occupational SOC codes. We include the following SOC codes: 15-1122 (Information Security Analysts), 15-1121 (Computer Systems Analysts), 15-1152 (Computer Network Support Specialists), 15-1141 (Database Administrators), 15-1142 (Network and Computer Systems Administrators), and 15-1143 (Computer Network Architects).

- **bgt_it_soc6_cyber_norm**

The normalized number of cybersecurity-related job postings posted by an entity, based on occupational SOC codes. We normalize by **bgt_soc6_total**, the total number of job postings posted by the entity.

- **bgt_it_soc6_cyber_rank**

The entity's cybersecurity rank compared to the other entities, based on occupational SOC codes.

- **bgt_it_soc6_cs_total**

The total number of computer-science-related job postings, based on occupational SOC codes. We include all SOC codes that begin with 15 (Computer and Mathematical Occupations).

- **bgt_it_soc6_cs_norm**

The normalized number of computer-science-related job postings posted by an entity, based on occupational SOC codes. We normalize by `bgt_soc6_total`, the total number of job postings posted by the entity.

- **`bgt_it_soc6_cs_rank`**

The entity's computer science rank compared to the other entities, based on occupational SOC codes.

2. Skill-Cluster-Families-based Variables

- **`bgt_scf_total`**

The total number of skill cluster families across job postings posted by an entity.

- **`bgt_it_scf_IT_total`**

The total number of skill cluster families across job postings posted by an entity that is related to the concept "Information Technologies". We include the following skill cluster family: "Information Technologies".

- **`bgt_it_scf_IT_norm`**

The normalized number of skill cluster families across job postings posted by an entity, that falls within the skill cluster family "Information Technology". We normalize by `bgt_scf_total`, the total number of skill cluster families across job postings posted by an entity.

- **`bgt_it_scf_IT_rank`**

The entity's Information Technology rank compared to the other entities, based on skill cluster families.

3. Skill-Clusters based Variables

- **`bgt_sc_total`**

The total number of skill clusters across job postings posted by an entity.

- **`bgt_it_sc_cloud_total`**

The total number of skill clusters across job postings posted by an entity that is related to the concept "Cloud". We include the following skill clusters: "Cloud Solutions", "Cloud Storage", "Cloud Computing", "Big Data", "Data Warehousing", "Database Management Systems", "NoSQL Databases", "Data Wrangling", "Graph Databases", and "Remote Desktop Software".

- **`bgt_it_sc_cloud_norm`**

The normalized number of skill clusters across job postings posted by an entity, that is related to the concept "Cloud". We normalize by `bgt_sc_total`, the total number of skill clusters across job postings posted by an entity.

- **`bgt_it_sc_cloud_rank`**

The entity's Cloud rank compared to the other entities, based on skill clusters.

- **bgt_it_sc_cyber_total**

The total number of skill clusters across job postings posted by an entity that is related to the concept “Cybersecurity”. We include the following skill clusters: “Cybersecurity”, “Network Security”, “Technical Support”, “Database Administration”, “Data Management”, “Information Security”, “Application Security”, and “Internet Security”.

- **bgt_it_sc_cyber_norm**

The normalized number of skill clusters across job postings posted by an entity, that is related to the concept “Cybersecurity”. We normalize by `bgt_sc_total`, the total number of skill clusters across job postings posted by an entity.

- **bgt_it_sc_cyber_rank**

The entity’s Cybersecurity rank compared to the other entities, based on skill clusters.