

National Flood Insurance Program (NFIP) Data Frequently Asked Questions (FAQs)

Q: Why can't the National Flood Insurance Program (NFIP) provide address-level data?

A: FEMA has a responsibility to protect policyholder privacy pursuant to the Privacy Act of 1974. In the data published in June 2019, FEMA provided the most granular data possible to promote transparency while protecting customer privacy consistent with the Privacy Act of 1974 and the Freedom of Information Act (FOIA). This is consistent with additional FEMA programmatic datasets posted on OpenFEMA as well.

Q: Why is the NFIP providing this data?

A: FEMA believes this historic release of NFIP data promotes transparency, reduces complexity related to public data requests, and improves how stakeholders interact with and understand the program.

This is the largest, most comprehensive release of NFIP data coordinated by FEMA to date. Academic institutions, the private sector, and the public are able to access and conduct customizable searches that allow creating reports and analyzing and visualizing present and historical NFIP data faster and easier than before. OpenFEMA demonstrates FEMA's commitment to building a culture of preparedness in the nation because it provides this important claims and policy information people need to make better choices about their flood risk and the insurance they need to protect the life they've built.

Q: How do I use the policy and claims datasets?

A: As part of the OpenFEMA initiative, FEMA is providing compressed zip files for download with the data in a .CSV format. In a subsequent release of the data there will be read-only, API-based access to flood insurance policies and claims datasets.

Q: Will the datasets contain historical data or is there a specific cutoff?

A: Anyone can use OpenFEMA to access the most current flood insurance claims and policies datasets. FEMA will post the program's full claim history with a date range from 1978 to present; and a policy dataset with 10-years' worth of policy records.

Q: How are you protecting policyholders' privacy?

A: Personal identifiable information (PII) is redacted and data is anonymized to the census tract, reported zip code, and to one decimal point (.1) digit of latitude and longitude. If mapped, flood insurance policies and claims may appear to be clustered at a particular location due to this anonymization.

Q: Why doesn't this data align with totals previously reported by the NFIP?

A: The NFIP legacy system operates on a monthly cycle and it takes several months to resolve flood insurance policies and claims errors. As a result, NFIP reporting is not static and totals will vary from one month to the next. FEMA has spent the past several years modernizing its reporting infrastructure and re-staging its legacy data in a modern database to improve both data quality and reporting operating standards.

Q: Who benefits from data sharing?

A: This data is now available for a number of public entities including: public institutions; state and local governments; Write Your Own (WYO) companies; the NFIP's Direct Servicing Agent (NFIP Direct); NFIP Vendors; academia; private insurers; capital markets; private sector industries; non-profit and foundations; and other federal government agencies.

Q: What has FEMA done to ensure that this data is correct and of high quality?

A: FEMA is reporting the data to the public as it was reported and recorded in the NFIP legacy system. The NFIP legacy system includes business rules which reject critical transaction errors but allow some non-critical errors in order to facilitate the reasonable availability of insurance given an antiquated technological platform.

To reduce confusion and ensure quality data, FEMA has omitted fields in the current dataset for which there is low fidelity. In addition, the NFIP has some policy rules that allow properties to be rated in one flood zone, while being physically located in another. Over time, the NFIP's recorded and reported geographical data has become more precise, so more recent data will be more reliable for geographical analysis than older data.

Q: Should this data be used in commercial catastrophic models to evaluate risk?

A: FEMA does not recommend the use of this data as input for commercial catastrophic models. The truncated coordinates do not provide an accurate depiction of spatial risk and will not result in credible model results. Moreover, the displacement of risks due to truncated coordinates will impact the appropriateness of the elevation data provided. FEMA recommends using the provided elevation data as a descriptive field for aggregation only. FEMA does not intend to release assumptions for mapping OpenFEMA data to various commercial catastrophe model inputs.

Q: Why is the policy dataset provided in segments?

A: In order to improve accessibility and performance we have truncated and compressed the policies dataset. The first segment includes headers and meta data, all additional segments do not. The policy dataset is segmented based on record counts and not region or time frame. If a user needs to extract a subset of the policy data supplied by FEMA, a guide provides some possible approaches for collating and extracting data from the full policy file.

Q: How often will the data be updated?

A: FEMA will update the data every 40 to 60 days.

Q: How do I get policies in-force for a given point in time?

A: The policy file is a transactional dataset which shows when a policy was in force and eventually terminated. In order to understand the policies or policy contracts in force at a point in time the user should only consider records with a “policyeffectivedate” on or before a given date and a “policyterminationdate” after that same date.

Q: How can I calculate the total claim payments made in a given year?

A: The current claims data does not include payment date or claims transaction. Users can use the date of loss to understand the magnitude of losses for a given calendar year but it is not possible to report disbursements by year from this dataset.

Q: Why is there latitude and longitude in this dataset, I thought the datasets were anonymized?

A: Latitude and longitude can be used to identify a specific location. In order to protect policy holder privacy, latitude and longitude are truncated to one decimal point. The more significant digits (decimals), the more precise a set of coordinates can be. The level of data that FEMA provides for users will find the coordinates accurate to within approximately six miles.

Q: Which geographical fields are best to use for aggregation?

A: Census tract and county fields are best to use for aggregation since they are derived from policy or claim geocode.