

scrapy Unit Tests

scrapy

“ An open source and collaborative framework for extracting the data you need from websites. In a fast, simple, yet extensible way. ”

- 43k GitHub stars
- Python 3 codebase

scrapy Unit Testing

- 52 (passing) unit tests
- 78 asserts

Test Suites

- Via `pytest`'s marks
- A pool of principles made by combining Right-BICEP and CORRECT

Mark	Count
crawlers_testing	1
offline	51
online	1
principle_cardinality_0	4
principle_cardinality_1	4
principle_cardinality_n	4
principle_conformance	9
principle_error	10
principle_existence	4
principle_inverse	10
principle_performance	52
principle_range_lower	12
principle_right	52
principle_time	52
robotstxt_testing	3
sitemap_testing	8
technique_fake	6
technique_monkey	5

Example Unit Test I

```
def gzip_wrapper(filename: str, data: bytes) -> int:
    with open(filename, "wb") as file:
        plugin = GzipPlugin(
            file,
            feed_options={
                "gzip_compresslevel": 9,
                "gzip_mtime": 0,
                "gzip_filename": "",
            },
        )
        return_code = plugin.write(data)
        plugin.close()

    return return_code
```

Example Unit Test II

```
@pytest.mark.principle_right
@pytest.mark.principle_inverse
@pytest.mark.principle_time
@pytest.mark.principle_performance
@pytest.mark.offline
@pytest.mark.timeout(0.1)
def test_gzip_valid_content() -> None:
    """Tests if the content is written correctly into the GZIP file."""
    data = b"data"

    with tempfile.NamedTemporaryFile("wb", delete=False) as temp:
        return_code = gzip_wrapper(temp.name, data)
        assert return_code != 0, "The written size for GZIP is zero."

    uncompressed_content = ungzip(temp.name)
    assert (
        data == uncompressed_content
    ), "The written data for GZIP is different from the original one."
```

Quality Controls

- Documentation for each file and function
- Suggestive message for asserts
- Auto-formatting with Black and isort
- Linting with Flake8
- Type checking with MyPy

Conclusions

- `scrapy` as SUT
- 52 unit tests
- Marks usage for suites separation
- Built-in quality controls