Name: Abon, Benedict Aldous A. Section: CPE22S3 Performed on: 02/19/2026 Submitted on: 02/19/2026 Submitted to: Engr. Neil Barton James Matira

Seatwork 7.2 Programming Exercise: Data Wrangling with Pandas - Part 2

Exercise Part 4:

1. Using the meteorite data from the Meteorite_Landings.csv file, create a pivot table that shows both the number of meteorites and the 95th percentile of meteorite mass for those that were found versus observed falling per year from 2005 through 2009 (inclusive). Hint: Be sure to convert the year column to a number as we did in the previous exercise.
2. Using the meteorite data from the Meteorite_Landings.csv file, compare summary statistics of the mass column for the meteorites that were found versus observed falling.

Exercise Part 5:

1. Using the taxi trip data in the 2019_Yellow_Taxi_Trip_Data.csv file, resample the data to an hourly frequency based on the dropoff time. Calculate the total trip_distance, fare_amount, tolls_amount, and tip_amount, then find the 5 hours with the most tips.

```
import pandas as pd
import numpy as np


A module that was compiled using NumPy 1.x cannot be run in
NumPy 2.4.2 as it may crash. To support both 1.x and 2.x
versions of NumPy, modules must be compiled with NumPy 2.0.
Some module may need to rebuild instead e.g. with 'pybind11>=2.12'.

If you are a user of the module, the easiest solution will be to
downgrade to 'numpy<2' or try to upgrade the affected module.
We expect that some modules will need time to support NumPy 2.

Traceback (most recent call last):  File "<frozen runpy>", line 198,
in _run_module_as_main
  File "<frozen runpy>", line 88, in _run_code
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\
ipykernel_launcher.py", line 17, in <module>
    app.launch_new_instance()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\traitlets\config\
application.py", line 1075, in launch_instance
    app.start()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelapp.py", line 701, in start
    self.io_loop.start()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\tornado\platform\
```

```
asyncio.py", line 205, in start
    self.asyncio_loop.run_forever()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\windows_events.py", line
322, in run_forever
    super().run_forever()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\base_events.py", line
641, in run_forever
    self._run_once()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\base_events.py", line
1986, in _run_once
    handle._run()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\events.py", line 88, in
_run
    self._context.run(self._callback, *self._args)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 534, in dispatch_queue
    await self.process_one()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 523, in process_one
    await dispatch(*args)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 429, in dispatch_shell
    await result
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 767, in execute_request
    reply_content = await reply_content
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
ipkernel.py", line 429, in do_execute
    res = shell.run_cell(
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
zmqshell.py", line 549, in run_cell
    return super().run_cell(*args, **kwargs)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3075, in run_cell
    result = self._run_cell(
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3130, in _run_cell
    result = runner(coro)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
async_helpers.py", line 128, in _pseudo_sync_runner
    coro.send(None)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3334, in run_cell_async
    has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3517, in run_ast_nodes
    if await self.run_code(code, result, async_=asy):
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3577, in run_code
```

```
    exec(code_obj, self.user_global_ns, self.user_ns)
  File "C:\Users\Aldous\AppData\Local\Temp\
ipykernel_3124\2162656668.py", line 1, in <module>
    import pandas as pd
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\
__init__.py", line 34, in <module>
    from pandas.compat import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\compat\
__init__.py", line 28, in <module>
    from pandas.compat.pyarrow import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\compat\
pyarrow.py", line 12, in <module>
    import pyarrow as pa
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pyarrow\
__init__.py", line 65, in <module>
    import pyarrow.lib as _lib

---------------------------------------------------------------------
-----
ImportError                               Traceback (most recent call
last)
File c:\Users\Aldous\anaconda3\Lib\site-packages\numpy\core\
_multiarray_umath.py:46, in __getattr__(attr_name)
     41     # Also print the message (with traceback).  This is
because old versions
     42     # of NumPy unfortunately set up the import to replace (and
hide) the
     43     # error.  The traceback shouldn't be needed, but e.g.
pytest plugins
     44     # seem to swallow it and we should be failing anyway...
     45     sys.stderr.write(msg + tb_msg)
---> 46     raise ImportError(msg)
     48 ret = getattr(_multiarray_umath, attr_name, None)
     49 if ret is None:

ImportError:
A module that was compiled using NumPy 1.x cannot be run in
NumPy 2.4.2 as it may crash. To support both 1.x and 2.x
versions of NumPy, modules must be compiled with NumPy 2.0.
Some module may need to rebuild instead e.g. with 'pybind11>=2.12'.

If you are a user of the module, the easiest solution will be to
downgrade to 'numpy<2' or try to upgrade the affected module.
We expect that some modules will need time to support NumPy 2.



A module that was compiled using NumPy 1.x cannot be run in
NumPy 2.4.2 as it may crash. To support both 1.x and 2.x
```

```
versions of NumPy, modules must be compiled with NumPy 2.0.
Some module may need to rebuild instead e.g. with 'pybind11>=2.12'.

If you are a user of the module, the easiest solution will be to
downgrade to 'numpy<2' or try to upgrade the affected module.
We expect that some modules will need time to support NumPy 2.

Traceback (most recent call last):  File "<frozen runpy>", line 198,
in _run_module_as_main
  File "<frozen runpy>", line 88, in _run_code
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\
ipykernel_launcher.py", line 17, in <module>
    app.launch_new_instance()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\traitlets\config\
application.py", line 1075, in launch_instance
    app.start()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelapp.py", line 701, in start
    self.io_loop.start()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\tornado\platform\
asyncio.py", line 205, in start
    self.asyncio_loop.run_forever()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\windows_events.py", line
322, in run_forever
    super().run_forever()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\base_events.py", line
641, in run_forever
    self._run_once()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\base_events.py", line
1986, in _run_once
    handle._run()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\events.py", line 88, in
_run
    self._context.run(self._callback, *self._args)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 534, in dispatch_queue
    await self.process_one()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 523, in process_one
    await dispatch(*args)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 429, in dispatch_shell
    await result
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 767, in execute_request
    reply_content = await reply_content
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
ipkernel.py", line 429, in do_execute
    res = shell.run_cell(
```

```
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
zmqshell.py", line 549, in run_cell
    return super().run_cell(*args, **kwargs)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3075, in run_cell
    result = self._run_cell(
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3130, in _run_cell
    result = runner(coro)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
async_helpers.py", line 128, in _pseudo_sync_runner
    coro.send(None)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3334, in run_cell_async
    has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3517, in run_ast_nodes
    if await self.run_code(code, result, async_=asy):
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3577, in run_code
    exec(code_obj, self.user_global_ns, self.user_ns)
  File "C:\Users\Aldous\AppData\Local\Temp\
ipykernel_3124\2162656668.py", line 1, in <module>
    import pandas as pd
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\
__init__.py", line 58, in <module>
    from pandas.core.api import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
api.py", line 9, in <module>
    from pandas.core.dtypes.dtypes import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
dtypes\dtypes.py", line 28, in <module>
    from pandas._libs import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pyarrow\
__init__.py", line 65, in <module>
    import pyarrow.lib as _lib

---------------------------------------------------------------------
-----
ImportError                               Traceback (most recent call
last)
File c:\Users\Aldous\anaconda3\Lib\site-packages\numpy\core\
_multiarray_umath.py:46, in __getattr__(attr_name)
     41     # Also print the message (with traceback).  This is
because old versions
     42     # of NumPy unfortunately set up the import to replace (and
hide) the
     43     # error.  The traceback shouldn't be needed, but e.g.
pytest plugins
     44     # seem to swallow it and we should be failing anyway...
```

```
     45     sys.stderr.write(msg + tb_msg)
---> 46     raise ImportError(msg)
     48 ret = getattr(_multiarray_umath, attr_name, None)
     49 if ret is None:

ImportError:
A module that was compiled using NumPy 1.x cannot be run in
NumPy 2.4.2 as it may crash. To support both 1.x and 2.x
versions of NumPy, modules must be compiled with NumPy 2.0.
Some module may need to rebuild instead e.g. with 'pybind11>=2.12'.

If you are a user of the module, the easiest solution will be to
downgrade to 'numpy<2' or try to upgrade the affected module.
We expect that some modules will need time to support NumPy 2.




A module that was compiled using NumPy 1.x cannot be run in
NumPy 2.4.2 as it may crash. To support both 1.x and 2.x
versions of NumPy, modules must be compiled with NumPy 2.0.
Some module may need to rebuild instead e.g. with 'pybind11>=2.12'.

If you are a user of the module, the easiest solution will be to
downgrade to 'numpy<2' or try to upgrade the affected module.
We expect that some modules will need time to support NumPy 2.

Traceback (most recent call last):  File "<frozen runpy>", line 198,
in _run_module_as_main
  File "<frozen runpy>", line 88, in _run_code
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\
ipykernel_launcher.py", line 17, in <module>
    app.launch_new_instance()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\traitlets\config\
application.py", line 1075, in launch_instance
    app.start()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelapp.py", line 701, in start
    self.io_loop.start()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\tornado\platform\
asyncio.py", line 205, in start
    self.asyncio_loop.run_forever()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\windows_events.py", line
322, in run_forever
    super().run_forever()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\base_events.py", line
641, in run_forever
    self._run_once()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\base_events.py", line
1986, in _run_once
```

```
    handle._run()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\events.py", line 88, in
_run
    self._context.run(self._callback, *self._args)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 534, in dispatch_queue
    await self.process_one()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 523, in process_one
    await dispatch(*args)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 429, in dispatch_shell
    await result
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 767, in execute_request
    reply_content = await reply_content
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
ipkernel.py", line 429, in do_execute
    res = shell.run_cell(
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
zmqshell.py", line 549, in run_cell
    return super().run_cell(*args, **kwargs)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3075, in run_cell
    result = self._run_cell(
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3130, in _run_cell
    result = runner(coro)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
async_helpers.py", line 128, in _pseudo_sync_runner
    coro.send(None)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3334, in run_cell_async
    has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3517, in run_ast_nodes
    if await self.run_code(code, result, async_=asy):
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3577, in run_code
    exec(code_obj, self.user_global_ns, self.user_ns)
  File "C:\Users\Aldous\AppData\Local\Temp\
ipykernel_3124\2162656668.py", line 1, in <module>
    import pandas as pd
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\
__init__.py", line 58, in <module>
    from pandas.core.api import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
api.py", line 27, in <module>
    from pandas.core.arrays import Categorical
```

```
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
arrays\__init__.py", line 1, in <module>
    from pandas.core.arrays.arrow import ArrowExtensionArray
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
arrays\arrow\__init__.py", line 5, in <module>
    from pandas.core.arrays.arrow.array import ArrowExtensionArray
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
arrays\arrow\array.py", line 65, in <module>
    from pandas.core import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\ops\
__init__.py", line 9, in <module>
    from pandas.core.ops.array_ops import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\ops\
array_ops.py", line 55, in <module>
    from pandas.core.computation import expressions
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
computation\expressions.py", line 22, in <module>
    from pandas.core.computation.check import NUMEXPR_INSTALLED
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
computation\check.py", line 5, in <module>
    ne = import_optional_dependency("numexpr", errors="warn")
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\compat\
_optional.py", line 158, in import_optional_dependency
    module = importlib.import_module(name)
  File "c:\Users\Aldous\anaconda3\Lib\importlib\__init__.py", line 90,
in import_module
    return _bootstrap._gcd_import(name[level:], package, level)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\numexpr\
__init__.py", line 24, in <module>
    from numexpr.interpreter import MAX_THREADS, use_vml,
__BLOCK_SIZE1__

---------------------------------------------------------------------
-----
AttributeError                            Traceback (most recent call
last)
AttributeError: _ARRAY_API not found


A module that was compiled using NumPy 1.x cannot be run in
NumPy 2.4.2 as it may crash. To support both 1.x and 2.x
versions of NumPy, modules must be compiled with NumPy 2.0.
Some module may need to rebuild instead e.g. with 'pybind11>=2.12'.

If you are a user of the module, the easiest solution will be to
downgrade to 'numpy<2' or try to upgrade the affected module.
We expect that some modules will need time to support NumPy 2.

Traceback (most recent call last):  File "<frozen runpy>", line 198,
in _run_module_as_main
```

```
  File "<frozen runpy>", line 88, in _run_code
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\
ipykernel_launcher.py", line 17, in <module>
    app.launch_new_instance()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\traitlets\config\
application.py", line 1075, in launch_instance
    app.start()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelapp.py", line 701, in start
    self.io_loop.start()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\tornado\platform\
asyncio.py", line 205, in start
    self.asyncio_loop.run_forever()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\windows_events.py", line
322, in run_forever
    super().run_forever()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\base_events.py", line
641, in run_forever
    self._run_once()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\base_events.py", line
1986, in _run_once
    handle._run()
  File "c:\Users\Aldous\anaconda3\Lib\asyncio\events.py", line 88, in
_run
    self._context.run(self._callback, *self._args)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 534, in dispatch_queue
    await self.process_one()
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 523, in process_one
    await dispatch(*args)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 429, in dispatch_shell
    await result
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
kernelbase.py", line 767, in execute_request
    reply_content = await reply_content
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
ipkernel.py", line 429, in do_execute
    res = shell.run_cell(
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\ipykernel\
zmqshell.py", line 549, in run_cell
    return super().run_cell(*args, **kwargs)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3075, in run_cell
    result = self._run_cell(
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3130, in _run_cell
    result = runner(coro)
```

```
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
async_helpers.py", line 128, in _pseudo_sync_runner
    coro.send(None)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3334, in run_cell_async
    has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3517, in run_ast_nodes
    if await self.run_code(code, result, async_=asy):
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\IPython\core\
interactiveshell.py", line 3577, in run_code
    exec(code_obj, self.user_global_ns, self.user_ns)
  File "C:\Users\Aldous\AppData\Local\Temp\
ipykernel_3124\2162656668.py", line 1, in <module>
    import pandas as pd
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\
__init__.py", line 58, in <module>
    from pandas.core.api import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
api.py", line 27, in <module>
    from pandas.core.arrays import Categorical
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
arrays\__init__.py", line 1, in <module>
    from pandas.core.arrays.arrow import ArrowExtensionArray
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
arrays\arrow\__init__.py", line 5, in <module>
    from pandas.core.arrays.arrow.array import ArrowExtensionArray
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
arrays\arrow\array.py", line 79, in <module>
    from pandas.core.arrays.masked import BaseMaskedArray
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
arrays\masked.py", line 56, in <module>
    from pandas.core import (
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\core\
nanops.py", line 54, in <module>
    bn = import_optional_dependency("bottleneck", errors="warn")
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\pandas\compat\
_optional.py", line 158, in import_optional_dependency
    module = importlib.import_module(name)
  File "c:\Users\Aldous\anaconda3\Lib\importlib\__init__.py", line 90,
in import_module
    return _bootstrap._gcd_import(name[level:], package, level)
  File "c:\Users\Aldous\anaconda3\Lib\site-packages\bottleneck\
__init__.py", line 7, in <module>
    from .move import (move_argmax, move_argmin, move_max, move_mean,
move_median,

---------------------------------------------------------------------
-----
ImportError                               Traceback (most recent call
```

```
last)
File c:\Users\Aldous\anaconda3\Lib\site-packages\numpy\core\
_multiarray_umath.py:46, in __getattr__(attr_name)
     41     # Also print the message (with traceback).  This is
because old versions
     42     # of NumPy unfortunately set up the import to replace (and
hide) the
     43     # error.  The traceback shouldn't be needed, but e.g.
pytest plugins
     44     # seem to swallow it and we should be failing anyway...
     45     sys.stderr.write(msg + tb_msg)
---> 46     raise ImportError(msg)
     48 ret = getattr(_multiarray_umath, attr_name, None)
     49 if ret is None:

ImportError:
A module that was compiled using NumPy 1.x cannot be run in
NumPy 2.4.2 as it may crash. To support both 1.x and 2.x
versions of NumPy, modules must be compiled with NumPy 2.0.
Some module may need to rebuild instead e.g. with 'pybind11>=2.12'.

If you are a user of the module, the easiest solution will be to
downgrade to 'numpy<2' or try to upgrade the affected module.
We expect that some modules will need time to support NumPy 2.
```

Exercise Part 4

```
# Meteorite data
meteor_landings = pd.read_csv('../[00]
Datasets/Meteorite_Landings.csv')
meteor_landings.head()
```

```
       name   id nametype     recclass  mass (g)  fall  \
0    Aachen    1    Valid           L5      21.0  Fell
1    Aarhus    2    Valid           H6     720.0  Fell
2      Abee    6    Valid         EH4  107000.0  Fell
3  Acapulco   10    Valid  Acapulcoite    1914.0  Fell
4   Achiras  370    Valid           L6     780.0  Fell

                      year    reclat    reclong           GeoLocation
0  01/01/1880 12:00:00 AM  50.77500    6.08333     (50.775, 6.08333)
1  01/01/1951 12:00:00 AM  56.18333   10.23333  (56.18333, 10.23333)
2  01/01/1952 12:00:00 AM  54.21667 -113.00000    (54.21667, -113.0)
3  01/01/1976 12:00:00 AM  16.88333  -99.90000     (16.88333, -99.9)
4  01/01/1902 12:00:00 AM -33.16667  -64.95000   (-33.16667, -64.95)
```

```
# Update the year column to only contain the year
meteor_landings['year'] = meteor_landings['year'].str.slice(6, 10)
```

```
meteor_landings['year'] = pd.to_numeric(meteor_landings['year'],
errors='coerce').astype('Int64')

meteor_landings.head()
```

```
       name   id nametype     recclass  mass (g)  fall  year    reclat
\
0    Aachen    1    Valid           L5      21.0  Fell  1880  50.77500

1    Aarhus    2    Valid           H6     720.0  Fell  1951  56.18333

2      Abee    6    Valid          EH4  107000.0  Fell  1952  54.21667

3  Acapulco   10    Valid  Acapulcoite    1914.0  Fell  1976  16.88333

4   Achiras  370    Valid           L6     780.0  Fell  1902 -33.16667


     reclong           GeoLocation
0    6.08333     (50.775, 6.08333)
1   10.23333  (56.18333, 10.23333)
2 -113.00000    (54.21667, -113.0)
3  -99.90000     (16.88333, -99.9)
4  -64.95000   (-33.16667, -64.95)
```

```
# Pivot table: meteorite count and 95th percentile mass by fall type
(2005-2009)
meteor_landings['mass (g)'] = pd.to_numeric(meteor_landings['mass
(g)'], errors='coerce')

filtered = meteor_landings[
    meteor_landings['year'].between(2005, 2009)
    & meteor_landings['fall'].isin(['Found', 'Fell'])
]

pivot_2005_2009 = pd.pivot_table(
    filtered,
    index='year',
    columns='fall',
    values='mass (g)',
    aggfunc=['count', lambda s: s.quantile(0.95)]
)

pivot_2005_2009.columns = [
    f"{'meteorite_count' if stat == 'count' else
'mass_95th_pct_g'}_{fall_type.lower()}"
    for stat, fall_type in pivot_2005_2009.columns
]

pivot_2005_2009
```

```
      meteorite_count_fell  meteorite_count_found
mass_95th_pct_g_fell  \
year

2005                   NaN                  874.0
NaN
2006                   5.0                 2450.0
25008.0
2007                   8.0                 1181.0
89675.0
2008                   9.0                  948.0
106000.0
2009                   5.0                 1492.0
8333.4

      mass_95th_pct_g_found
year
2005                4500.00
2006                1600.50
2007                1126.90
2008                2274.80
2009                1397.25
```

```
# found meteor masses
found_meteorites = meteor_landings[meteor_landings['fall']=='Found']
found_meteorites = found_meteorites.sort_index()
found_meteorites
```

```
                        name     id nametype               recclass
mass (g)  \
37     Northwest Africa 5815  50693    Valid                     L5
256.8
520      Cumulus Hills 04075  32531    Valid             Pallasite
9.6
757     Dominion Range 03239  32591    Valid                     L6
69.5
804     Dominion Range 03240  32592    Valid                   LL5
290.9
1111                  Abajo      4    Valid                     H5
331.0
...                     ...    ...      ...                    ...
...
45711             Zillah 002  31356    Valid                Eucrite
172.0
45712                 Zinder  30409    Valid  Pallasite, ungrouped
46.0
45713                   Zlin  30410    Valid                     H4
3.3
45714              Zubkovsky  31357    Valid                     L6
2167.0
```

```
45715             Zulu Queen  30414    Valid                   L3.7
200.0

        fall  year    reclat    reclong             GeoLocation
37     Found  <NA>   0.00000    0.00000             (0.0, 0.0)
520    Found  2003       NaN        NaN                    NaN
757    Found  2002       NaN        NaN                    NaN
804    Found  2002       NaN        NaN                    NaN
1111   Found  1982  26.80000 -105.41667     (26.8, -105.41667)
...      ...   ...       ...        ...                    ...
45711  Found  1990  29.03700   17.01850      (29.037, 17.0185)
45712  Found  1999  13.78333    8.96667    (13.78333, 8.96667)
45713  Found  1939  49.25000   17.66667      (49.25, 17.66667)
45714  Found  2003  49.78917   41.50460    (49.78917, 41.5046)
45715  Found  1976  33.98333 -115.68333  (33.98333, -115.68333)

[44609 rows x 10 columns]
```

```
fallen_meteorites = meteor_landings[meteor_landings['fall']=='Fell']
fallen_meteorites.sort_index()
fallen_meteorites
```

```
          name     id nametype     recclass  mass (g)  fall  year
reclat  \
0       Aachen      1    Valid           L5      21.0  Fell  1880
50.77500
1       Aarhus      2    Valid           H6     720.0  Fell  1951
56.18333
2         Abee      6    Valid          EH4  107000.0  Fell  1952
54.21667
3     Acapulco     10    Valid  Acapulcoite    1914.0  Fell  1976
16.88333
4      Achiras    370    Valid           L6     780.0  Fell  1902 -
33.16667
...        ...    ...      ...          ...       ...   ...   ...
...
1106  Zhuanghe  30408    Valid           H5    2900.0  Fell  1976
39.66667
1107     Zmenj  30411    Valid    Howardite     246.0  Fell  1858
51.83333
1108     Zomba  30412    Valid           L6    7500.0  Fell  1899 -
15.18333
1109   Zsadany  30413    Valid           H5     552.0  Fell  1875
46.93333
1110   Zvonkov  30415    Valid           H6    2568.0  Fell  1955
50.20000

        reclong            GeoLocation
0       6.08333      (50.775, 6.08333)
1      10.23333   (56.18333, 10.23333)
```

```
2    -113.00000     (54.21667, -113.0)
3     -99.90000      (16.88333, -99.9)
4     -64.95000    (-33.16667, -64.95)
...         ...                    ...
1106  122.98333  (39.66667, 122.98333)
1107   26.83333   (51.83333, 26.83333)
1108   35.28333  (-15.18333, 35.28333)
1109   21.50000       (46.93333, 21.5)
1110   30.25000          (50.2, 30.25)

[1107 rows x 10 columns]
```

```
found_meteor_mean = np.mean(found_meteorites['mass (g)'])
found_meteor_median = np.median(found_meteorites['mass (g)'])
found_meteor_mode = np.array(found_meteorites['mass (g)'].mode())
found_meteor_var = np.var(found_meteorites['mass (g)'])
found_meteor_std = np.std(found_meteorites['mass (g)'])

fallen_meteor_mean = np.mean(fallen_meteorites['mass (g)'])
fallen_meteor_median = np.median(fallen_meteorites['mass (g)'])
fallen_meteor_mode = np.array(fallen_meteorites['mass (g)'].mode())
fallen_meteor_var = np.var(fallen_meteorites['mass (g)'])
fallen_meteor_std = np.std(fallen_meteorites['mass (g)'])

meteorites_mass_stats = pd.DataFrame({
    'fall_type': ['Found', 'Fell'],
    'mean_mass': [found_meteor_mean, fallen_meteor_mean],
    'median_mass': [found_meteor_median, fallen_meteor_median],
    'mode_mass': [found_meteor_mode, fallen_meteor_mode],
    'variance_mass': [found_meteor_var, fallen_meteor_var],
    'std_mass': [found_meteor_std, fallen_meteor_std]
})

meteorites_mass_stats
```

```
  fall_type     mean_mass  median_mass                  mode_mass  \
0     Found  12461.922983          NaN                      [1.3]
1      Fell  47070.715023          NaN  [1000.0, 2000.0, 4000.0]

   variance_mass       std_mass
0   3.261545e+11  571099.336798
1   5.137070e+11  716733.528633
```

Exercise Part 5

```
# Taxi data: hourly totals and top 5 hours by tip amount
taxi = pd.read_csv('../[00] Datasets/2019_Yellow_Taxi_Trip_Data.csv')

dropoff_col_candidates = ['tpep_dropoff_datetime', 'dropoff_datetime',
'lpep_dropoff_datetime']
dropoff_col = next((c for c in dropoff_col_candidates if c in
```

```
taxi.columns), None)
if dropoff_col is None:
    raise KeyError('No dropoff datetime column found in taxi
dataset.')
taxi[dropoff_col] = pd.to_datetime(taxi[dropoff_col], errors='coerce')

sum_cols = ['trip_distance', 'fare_amount', 'tolls_amount',
'tip_amount']
for col in sum_cols:
    taxi[col] = pd.to_numeric(taxi[col], errors='coerce')

hourly_totals = (
    taxi.dropna(subset=[dropoff_col])
        .set_index(dropoff_col)
        .resample('h')[sum_cols]
        .sum()
        .sort_index()
 )

top_5_tip_hours = hourly_totals.nlargest(5, 'tip_amount')
top_5_tip_hours
```

tpep_dropoff_datetime	trip_distance	fare_amount	tolls_amount	tip_amount
2019-10-23 16:00:00	10676.95	67797.76	699.04	12228.64
2019-10-23 17:00:00	16052.83	70131.91	4044.04	12044.03
2019-10-23 18:00:00	3104.56	11565.56	1454.67	1907.64
2019-10-23 15:00:00	14.34	213.50	0.00	51.75
2019-10-23 19:00:00	98.59	268.00	24.48	25.74

End