

git_comments:

1. TODO(BEAM-10570): Consider a pythonsdk logical type.
2. Before PEP 468, these are randomly ordered. At least provide deterministic behavior here. pylint: disable=dict-items-not-iterating
3. type: typing.Union[str, callable] type: typing.Union[callable, CombineFn] type: str
4. type: typing.Union[str, callable] type: typing.Union[str, callable]
5. pylint: disable=dict-items-not-iterating
6. [0], [-1, 1] [-2] [2]

git_commits:

1. **summary:** Add relational GroupBy transform to Python. (#11963)
message: Add relational GroupBy transform to Python. (#11963)

github_issues:**github_issues_comments:****github_pulls:**

1. **title:** Add relational GroupBy transform to Python.
body: ----- Thank you for your contribution! Follow this checklist to help us incorporate your contribution quickly and easily: - [] **Choose reviewer(s)** (https://beam.apache.org/contribute/#make-your-change) and mention them in a comment (R: @username). - [] **Format the pull request title** like '[BEAM-XXX] Fixes bug in ApproximateQuantiles', where you replace 'BEAM-XXX' with the appropriate JIRA issue, if applicable. This will automatically link the pull request to the issue. - [] **Update 'CHANGES.md'** with noteworthy changes. - [] **If this contribution is large, please file an Apache [Individual Contributor License Agreement]** (https://www.apache.org/licenses/icla.pdf). See the [Contributor Guide](https://beam.apache.org/contribute) for more tips on [how to make review process smoother](https://beam.apache.org/contribute/#make-reviewers-job-easier). Post-Commit Tests Status (on master branch) ----- Lang | SDK | Apex | Dataflow | Flink | Gearpump | Samza |
Spark --- | --- | --- | --- | --- | --- | --- Go | 
(https://builds.apache.org/job/beam_PostCommit_Go/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Go/lastCompletedBuild/) | --- | --- | 
(https://builds.apache.org/job/beam_PostCommit_Go_VR_Flink/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Go_VR_Flink/lastCompletedBuild/) | --- | --- | 
(https://builds.apache.org/job/beam_PostCommit_Go_VR_Spark/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Go_VR_Spark/lastCompletedBuild/) Java | 
(https://builds.apache.org/job/beam_PostCommit_Java/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java/lastCompletedBuild/) | 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Apex/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Apex/lastCompletedBuild/) | 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Dataflow/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Dataflow/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Dataflow_Java11/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Dataflow_Java11/lastCompletedBuild/) | 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Flink/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Flink/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Flink_Java11/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Flink_Java11/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Java_PVR_Flink_Batch/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_PVR_Flink_Batch/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Java_PVR_Flink_Streaming/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_PVR_Flink_Streaming/lastCompletedBuild/) | 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Gearpump/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Gearpump/lastCompletedBuild/) | 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Samza/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Samza/lastCompletedBuild/) | 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Spark/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_Spark/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Java_PVR_Spark_Batch/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_PVR_Spark_Batch/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_SparkStructuredStreaming/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Java_ValidatesRunner_SparkStructuredStreaming/lastCompletedBuild/) Python | 
(https://builds.apache.org/job/beam_PostCommit_Python2/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Python2/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Python35/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Python35/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Python36/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Python36/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Python37/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Python37/lastCompletedBuild/) | --- | 
(https://builds.apache.org/job/beam_PostCommit_Py_VR_Dataflow/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Py_VR_Dataflow/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Py_VR_Dataflow_V2/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Py_VR_Dataflow_V2/lastCompletedBuild/)
 
(https://builds.apache.org/job/beam_PostCommit_Py_ValCont/lastCompletedBuild/badge/icon))
(https://builds.apache.org/job/beam_PostCommit_Py_ValCont/lastCompletedBuild/) | 

(https://builds.apache.org/job/beam_PreCommit_Python2_PVR_Flink_Cron/lastCompletedBuild/badge/icon)
(https://builds.apache.org/job/beam_PreCommit_Python2_PVR_Flink_Cron/lastCompletedBuild/)
[![Build Status]
(https://builds.apache.org/job/beam_PostCommit_Python35_VR_Flink/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PostCommit_Python35_VR_Flink/lastCompletedBuild/) | --- | --- | [![Build Status]
(https://builds.apache.org/job/beam_PostCommit_Python_VR_Spark/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PostCommit_Python_VR_Spark/lastCompletedBuild/) XLang | --- | --- | [![Build Status]
(https://builds.apache.org/job/beam_PostCommit_XVR_Flink/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PostCommit_XVR_Flink/lastCompletedBuild/) | --- | --- | [![Build Status]
(https://builds.apache.org/job/beam_PostCommit_XVR_Spark/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PostCommit_XVR_Spark/lastCompletedBuild/) Pre-Commit Tests Status (on master branch) -----
----- | Java | Python | Go | Website --- | --- | --- | --- | --- Non-portable | [!
[Build Status](https://builds.apache.org/job/beam_PreCommit_Java_Cron/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PreCommit_Java_Cron/lastCompletedBuild/) | [![Build Status]
(https://builds.apache.org/job/beam_PreCommit_Python_Cron/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PreCommit_Python_Cron/lastCompletedBuild/)
[![Build Status]
(https://builds.apache.org/job/beam_PreCommit_PythonLint_Cron/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PreCommit_PythonLint_Cron/lastCompletedBuild/) | [![Build Status]
(https://builds.apache.org/job/beam_PreCommit_Go_Cron/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PreCommit_Go_Cron/lastCompletedBuild/) | [![Build Status]
(https://builds.apache.org/job/beam_PreCommit_Website_Cron/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PreCommit_Website_Cron/lastCompletedBuild/) Portable | --- | [![Build Status]
(https://builds.apache.org/job/beam_PreCommit_Portable_Python_Cron/lastCompletedBuild/badge/icon)]
(https://builds.apache.org/job/beam_PreCommit_Portable_Python_Cron/lastCompletedBuild/) | --- | --- See [.test-infra/jenkins/README]
(<https://github.com/apache/beam/blob/master/.test-infra/jenkins/README.md>) for trigger phrase, status and link of all Jenkins jobs.

github_pulls_comments:

1. CC: @TheNeuralBit
2. R: @TheNeuralBit This is ready for review.
3. I've refactored this to use named tuples rather than rows. PTAL.
4. Thanks. Filed a bug and added a TODO about a Python SDK logical type. I'll look at the inference separately.
5. Run Python PreCommit
6. :sdks:python:test-suites:dataflow:py2:preCommitIT_streaming is unrelated.

github_pulls_reviews:

1. ```suggestion beam.Map(lambda v: (expr(v), v)) | beam.GroupByKey() ```
2. Is this a fallback for when a Row uses types that we don't support in Python schemas? I worry about this since it makes it tricky for a user to tell when a PCollection can be used in an ExternalTransform that uses rows. Are there some specific types that we need to add coverage for?
3. I had this problem with the NamedTuple used in an external transform's output PCollection, I solved it there by making a custom `__reduce__` function that rebuilds from the proto-encoded schema:
https://github.com/apache/beam/blob/80de47655c90594eead2f51eece206e8c4fe0227/sdks/python/apache_beam/typehints/schemas.py#L210-L225 (A TODO is fine as well, just wanted to point that out in case it's helpful)
4. ```suggestion # TODO(BEAM-7372): Use {**a, **b} syntax once Python 2 is gone. ```
5. Will subclass actually raise an AttributeError? Curious what causes this
6. I'm pretty sure I did see this. (Maybe it was missing `__mro__` or something?)
7. Well, it's also there to handle the case where we simply don't know what the type is. I'll see if I can re-work this to use named tuples (thanks for the idea about `__reduce__`).
8. ```suggestion but provides several conveniences, e.g. ```
9. Ah I see. I'd rather not have this escape hatch if we can avoid it, but I can imagine it would be frustrating for users who don't care about using Rows xlang if we just refuse to make a RowCoder for them. Perhaps we should make `named_fields_to_schema` wrap unknown type hints in a "pythonsdk" logical type that uses fast primitives coder. Then we can give a better error message if/when a xlang issue arises. If you just add a TODO here I can take that action. When running locally I noticed that this is only needed for your tests because we can't infer a type for `sign=x // abs(x) if x else 0` and fallback to `Any`. Just curious - is it possible to fix the inference for that function? It seems like we should be able to.

jira_issues:

jira_issues_comments: