0.0.1 (DataHaskell)

Report by:	(Marco Zocca)
Participants:	(Nikita Tchayka, John Vial, ⋯)
Status:	(Alive and well)

The DataHaskell community was initiated in September 2016 as a gathering place for scientific computing, machine learning and data science practitioners and Haskell programmers; we observe a growing interest in using functional composition, domain-specific languages and type inference for implementing robust and reusable data processing pipelines.

DataHaskell revolves around a Gitter chatroom [1] and a GitHub organization [2]. The development team uses a Trello board to track the details of ongoing activities [3]; access to this tool will be granted to all interested parties.

One of our first steps was setting up a wiki [4] to serve as a knowledge base of related Haskell packages and frameworks.

After an informal survey we concluded that large part of our userbase seems to be lacking most

- an IDE for exploratory data analysis,
- a generic 'data-frame' for fast import and manipulation of heterogeneous tabular data,
- a native numerical back-end;

therefore current DataHaskell activities are first focusing on improving the ergonomics of the IHaskell notebook [5], and putting it to use on a Kaggle classification exercise [6]. This will serve to highlight the merits and the gaps or inefficiencies in the current package landscape.

We cherish the open and multidisciplinary nature of our community, and welcome all new users and contributions.

Further reading

datahaskell.org

[1] https://gitter.im/dataHaskell/Lobby
[2] https://github.com/DataHaskell
[3] https://trello.com/b/ucB25d5v/tasks
[4] http://wiki.datahaskell.org
[5] https://github.com/DataHaskell/DataIHaskell
[6] https://github.com/johnny555/ToolExamples/tree/master/Kaggle/Titanic