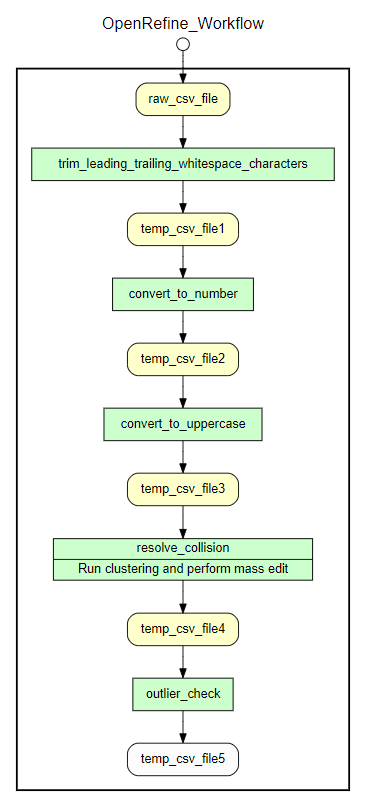
# @begin OpenRefine\_Workflow @desc Display one or more greetings to the user.

# @in raw\_csv\_file

# @begin trim\_leading\_trailing\_whitespace\_characters

# @in raw @as raw\_csv\_file

# @out pipe1out @as temp\_csv\_file1

# @end trim\_leading\_trailing\_whitespace\_characters

# @begin convert\_to\_number

# @in pipe1out @as temp\_csv\_file1

# @out pipe2out @as temp\_csv\_file2

# @end convert\_to\_number

# @begin convert\_to\_uppercase

# @in pipe2out @as temp\_csv\_file2

# @out pipe3out @as temp\_csv\_file3

# @end convert\_to\_uppercase

# @begin resolve\_collision @desc Run clustering and perform mass edit

# @in pipe3out @as temp\_csv\_file3

# @out pipe4out @as temp\_csv\_file4

# @end resolve\_collision

# @begin outlier\_check

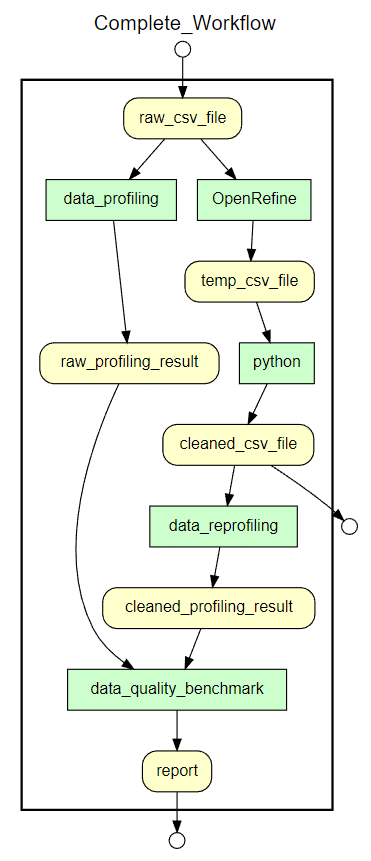
# @in pipe4out @as temp\_csv\_file4

# @out pipe5out @as temp\_csv\_file5

# @end outlier\_check

# @end OpenRefine\_Workflow

# @begin Complete\_Workflow @desc Display one or more greetings to the user.

# @in raw\_csv\_file

# @begin data\_profiling

# @in raw @as raw\_csv\_file

# @out raw\_profiling\_result

# @end data\_profiling

# @begin OpenRefine

# @in raw\_csv\_file

# @out temp\_csv\_file

# @end OpenRefine

# @begin python

# @in temp\_csv\_file

# @out cleaned\_csv\_file

# @end python

# @begin data\_reprofiling

# @in cleaned\_csv\_file

# @out cleaned\_profiling\_result

# @end data\_reprofiling

# @begin data\_quality\_benchmark

# @in cleaned\_profiling\_result

# @in raw\_profiling\_result

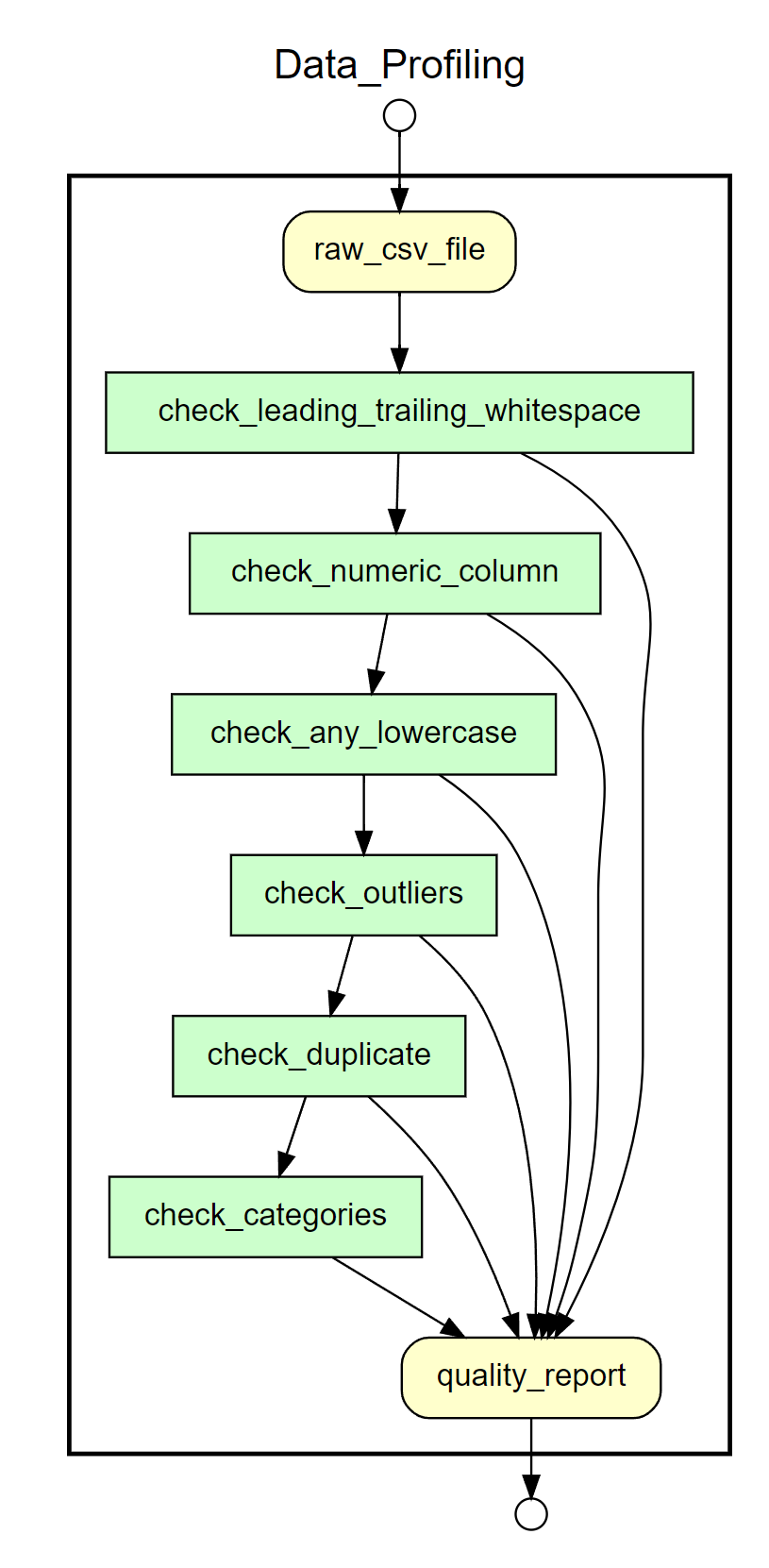
# @out report

# @end data\_quality\_benchmark

# @out report

# @out cleaned\_csv\_file

# @end Complete\_Workflow

# @begin Data\_Profiling

# @in raw\_csv\_file

# @begin check\_leading\_trailing\_whitespace

# @in raw @as raw\_csv\_file

# @out quality\_report

# @end check\_leading\_trailing\_whitespace

# @begin check\_numeric\_column

# @in check\_leading\_trailing\_whitespace

# @out quality\_report

# @end check\_numeric\_column

# @begin check\_any\_lowercase

# @in check\_numeric\_column

# @out quality\_report

# @end check\_any\_lowercase

# @begin check\_outliers

# @in check\_any\_lowercase

# @out quality\_report

# @end check\_outliers

# @begin check\_duplicate

# @in check\_outliers

# @out quality\_report

# @end check\_duplicate

# @begin check\_categories

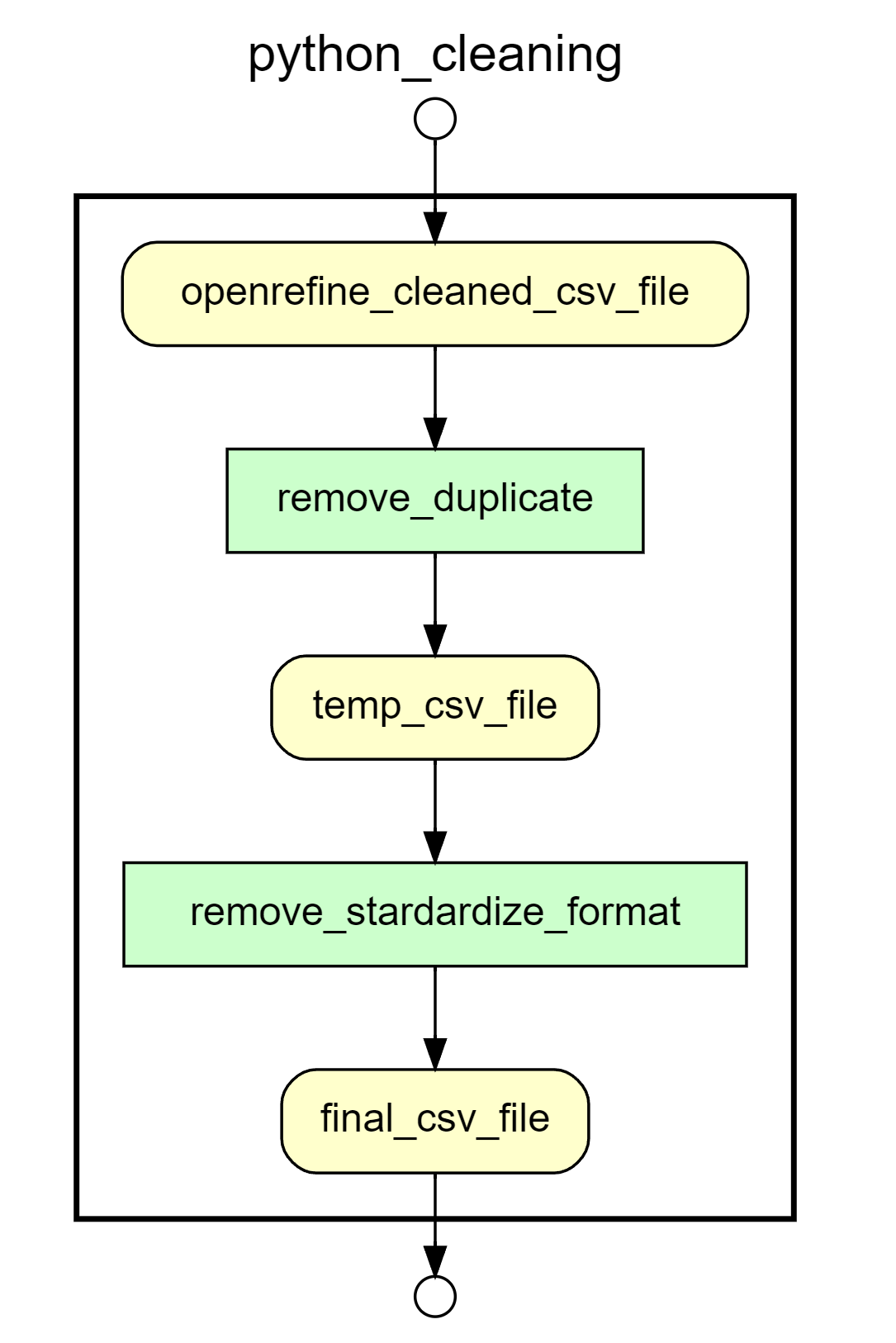
# @in check\_duplicate

# @out quality\_report

# @end check\_categories

# @out quality\_report

# @end OpenRefine\_Workflow

# @begin python\_cleaning

# @in openrefine\_cleaned\_csv\_file

# @begin remove\_duplicate

# @in openrefine\_cleaned\_csv\_file

# @out pipe1out @as temp\_csv\_file

# @end remove\_duplicate

# @begin remove\_stardardize\_format

# @in pipe1out @as temp\_csv\_file

# @out final\_csv\_file

# @end remove\_stardardize\_format

# @out final\_csv\_file

# @end OpenRefine\_Workflow