PharmaSUG 2022 - Paper QT-093

SAS® Studio: Creating, Analyzing, and Reporting Data with Built-in & Custom Tasks

Himanshu Patel, Chintan Pandya, Merck & Co., Inc., Rahway, NJ, USA

ABSTRACT

SAS® Enterprise Guide or Display Manager is the default programming tool for SAS users in the clinical trial industry. SAS Studio is a relatively new tool within the SAS environment, with powerful enhancements compared to SAS Enterprise Guide and Display Manager. SAS Studio provides built-in tasks for generating and executing complex statistical methods, models, graphs, and SAS procedures using a simple point-and-click interface. SAS Studio also allows programmers to develop custom tasks for the unique requirements they receive from time to time. This paper helps understand some of the predefined point-and-click task features and how to create custom tasks for individual needs with the SAS Studio interface. It can be a beneficial tool for urgent requests, data analysis, and exploration. This paper will be helpful to those planning to consider an alternative programming option to SAS EG or Display Manager.

INTRODUCTION

SAS users in the pharmaceutical industry have favored using SAS Enterprise Guide or Display Manager as a programming environment for more than 30 years now. In recent years, there has been a significant change in the computing platform, power, and how the new generation interacts with computing tools. As cloud-based computing presence increased in all organizations, the SAS applications also changed. The first version of SAS Studio 3.1 was released in March 2014, and lots of enhancements have been made in the latest version since then. It is a web-based interface and HTML5 application for SAS programmers. It is designed to support cloud-based or on-premise deployment. SAS Studio supports multiple browsers, such as Microsoft Edge, Safari, Mozilla Firefox, and Google Chrome.

In addition to writing and executing your own SAS programs, users can use the built-in point-and-click tasks that come with SAS Studio to analyze and explore data. The tasks are based on the SAS system procedures and provide access to the most commonly used graphs, statistical, and analytical procedures. Users can create custom tasks in SAS Studio by modifying existing tasks or developing new ones using basic text commands. The developed custom tasks can be used repeatedly to generate output with different parameters. This paper demonstrates how SAS Studio's adoption offers a familiar yet modern and improved programming environment. In this paper, we will discuss two important features of SAS Studio and how to use it to create, analyze and report data quickly and easily.

TASK BASICS

In SAS Studio, all tasks are based on the CTM (Common Task Model), which is made up of XML (Extensible Markup Language - similar to HTML, but without predefined tags to use), and VTL (Velocity Template Language). The CTM file defines the templates for the individual task and contains many different elements required for the user interface to generate the necessary SAS code. The CTM file contains many different elements.

The CTM file which generates the user interface has the "TASK" element, which has the following "CHILD" elements:

Element Name	Functionality
Registration	Defines the task type (e.g., task name, description, and properties)
Metadata	Defines whether the input data source is required. It also specifies any role and options in the task. • **Roles Element:** Variables types required by the task • **Options Element:** How to display option in the interface
User Interface (UI)	Describe the presentation of the interface to the user

Element Name	Functionality
Dependencies	Describe any dependencies on one another
Requirements	Specifies required conditions for code to be generated
Code Template	Determine the final output

Table 1. CHILD elements of the CTM (Common Task Model) and its functionality.

POINT-AND-CLICK TASKS

As discussed above, the point-and-click user interface offered by SAS Studio is another valuable feature for the users. This interface guides the user in creating, analyzing, reporting, and data conversion processes. The SAS code is generated simultaneously when a user selects the task and defines the required parameters (e.g., dataset, variables, options). The user simply needs to execute or run the task-generated code to get the required output.

The following table represents some of the important and frequently used built-in point-and-click tasks in SAS Studio.

Task Category	Task Name	SAS Procedures Used
Data	List Table Attributes	DATASETS
	Characterize Data	FREQ
	Describe Missing Data	FORMAT, FREQ
	List Data	PRINT
	Transpose Data	TRANSPOSE
	Stack/Split Columns	SORT, TRANSPOSE, DELETE
	Filter Data	SQL
	Sort Data	SORT
	Combine Tables	SQL
Graph	Bar Chart	SGPLOT
	Bar-Line Chart	MEANS, SGPLOT, DATASETS
	Box Plot	SGPLOT
	Scatter Plot	SGPLOT
	Pie Chart	TEMPLATE, SGRENDER
Statistics	Summary Statistics	MEANS
	One-Way Frequencies	FREQ
	Data Exploration	**Box Plot:** SORT, BOXPLOT, DATASETS
		Histogram: UNIVARIATE, SQL
Linear Models	One-Way ANOVA	GLM
	Linear Regression	GLMSELECT, REG, DELETE
	Mixed Models	MIXED
Survival Analysis	Nonparametric Survival Analysis	LIFETEST
	Proportional Hazards Regression	PHREG

Table 2. Build-in Task and SAS Procedures Associated with the Individual Task.

BAR CHART WITH POINT-AND-CLICK TASK

As discussed in Table 2, there are many built-in tasks for which SAS Studio provides a simple and guided user interface to get work done in a few clicks. As a result, it gives a clean and controllable environment for any experience level users to execute basic or complex SAS procedures quickly and efficiently.

The following screenshot shows on creating a bar chart with point-and-click functionality:

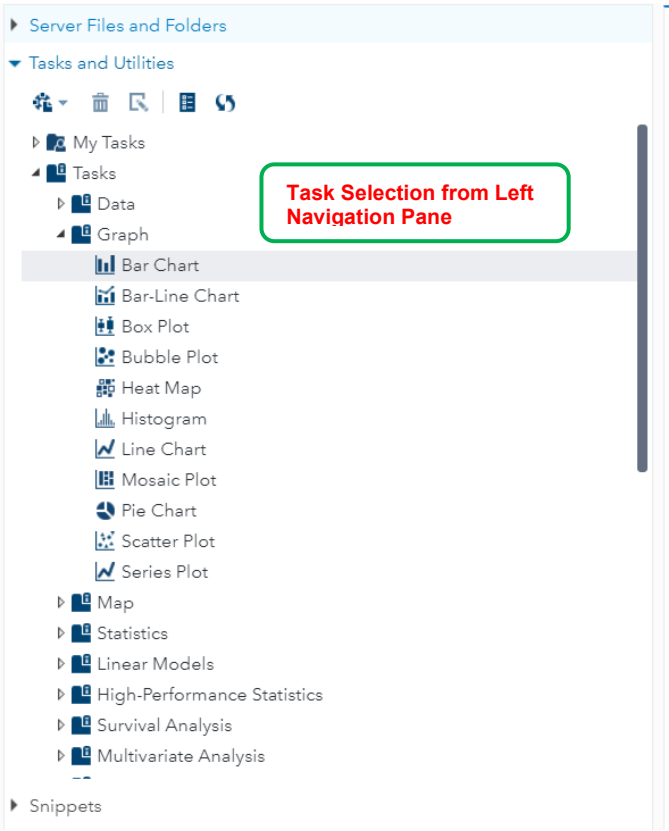

Figure 1. Select a task (Tasks and Utilities → Tasks → Graph → Bar Chart).

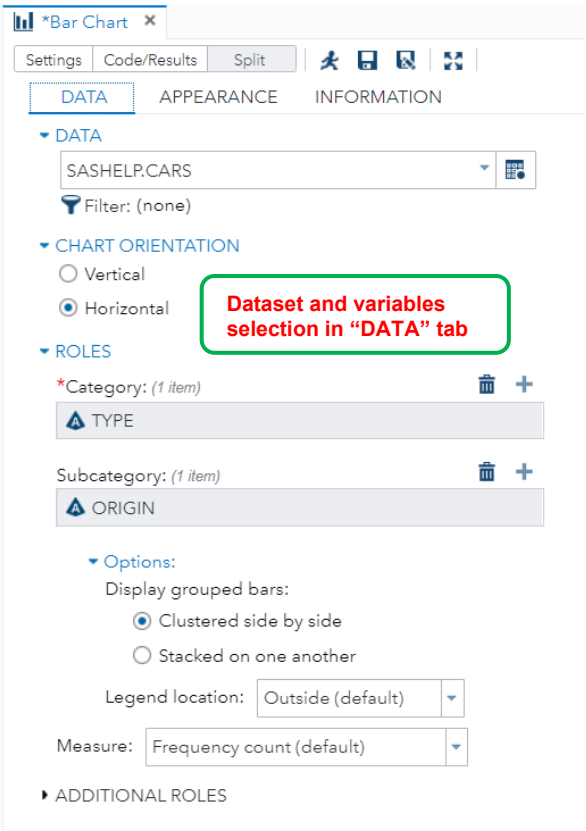

Figure 2. Filled out required elements under the "DATA" tab.

In figure 2 above, there are multiple options to select under the "DATA" tab. The user must choose all the required elements to generate the code. The optional elements can be selected as per the requirements or left empty.

- **DATA:** Select a dataset name from which plot needs to be generated *(required)*
- **CHART ORIENTATION:** Choose the layout of the plot *(required, default: vertical)*
- **ROLES:** Choose a variable(s) that needs to be displayed on the plot
  - **Category:** Primary variable name *(required)*
  - **Subcategory:** Secondary variable name for subcategory analysis *(optional)*
    - Options will populate when the variable is specified in the subcategory field. It gives the flexibility to the users to display the grouped bar side-by-side with primary variable or stacked on it
- **ADDITIONAL ROLES:** It provides an option to display a separate plot for each value present in a group by variable or to add weight to the plot *(optional)*
  - **Group Analysis by:** Name of the group by variable
  - **Weight:** Name of the weight variable

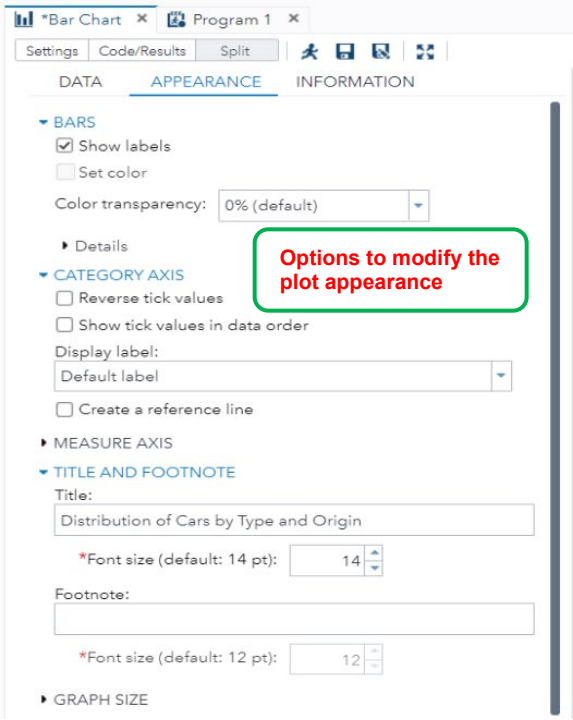

Figure 3. Set options to modify the plot appearance, including title and footnote.

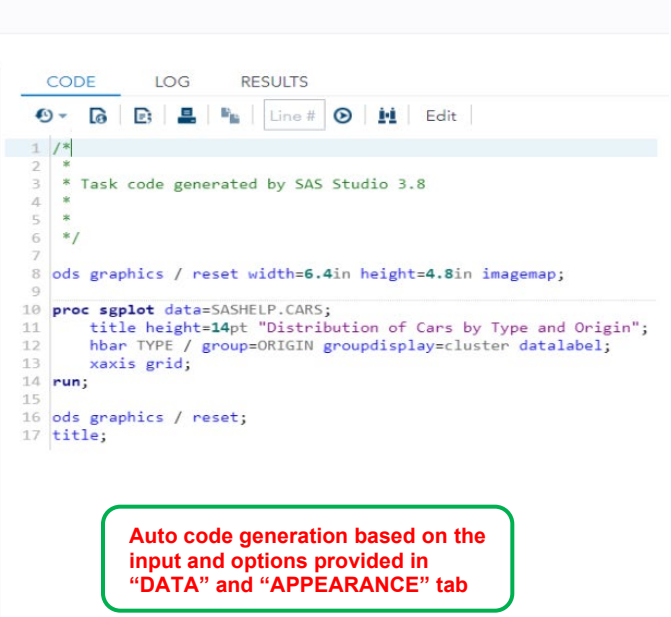

Figure 4. Auto-generated code for Box Plot based on the inputs in Figure 2 and Figure 3.

The following output is generated based on the execution of SAS code generated by the built-in task "Bar Chart" displayed in figure 4 above.

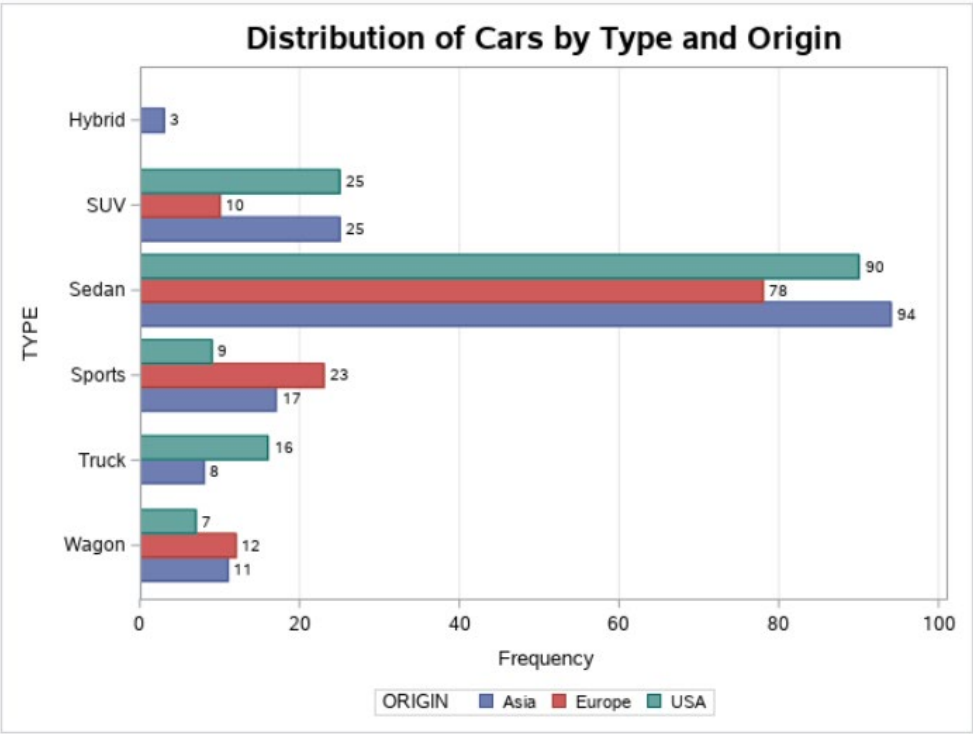

Figure 5. Built-in "Bar Chart" task applied to SASHELP.CARS and categorized by "Type" and "Origin".

CUSTOM TASK

There are times when the built-in task does not fit the requirements, and the user needs some specific type of analysis to generate the desired output. In this situation, developing custom tasks as per need can be a problem solver, but how difficult is it? Creating a custom task became more accessible, and there are multiple ways to develop a custom task.

1. **Modification to the existing predefined task**

   It is the easiest and quick way to create a custom task. This can be done by following the steps below:

   - Right-click on the built-in task and select the “Add to my Tasks” option
   - Provide a task name, description, category, and click “Add” in a pop-up window. The new task will be created and displayed under “My Tasks”
   - Right-click on the task name and choose the “edit” option. The task template code will open in a new editor
   - Review and modify the template as needed to generate the desired output

2. **Creating a new task from the template**

   There are three options to create a custom task from the template as described below:

   - New Task – It gives the user a blank template to create a custom task
   - Sample Task – It provides a basic task outline and controls available to the user
   - Advanced Task – It gives more complex functionalities to the user

CUSTOM TASK BASICS

Several online guides provide step-by-step instructions on creating custom tasks from scratch. It is a helpful resource for programmers of any experience level. Following are some basic steps to create a first custom task from scratch using the template described above.

1. Open the “New Task” template
2. Register a task
3. Specify data source and identify roles
4. Create options
5. Add Apache Velocity (AV) code
6. Run generated code to get the output

```
<?xml version="1.0" encoding="UTF-8"?>
<Task schemaVersion="5.4" runNLS="never">
        <Registration>
        <Name>New Task</Name>
        <Description>This is a blank task</Description>
        <GUID>726B2B12-4542-4200-96F4-960C645D767F</GUID>
        <Procedures>TBD</Procedures>
        <Version>3.8</Version>
        <Links>
                <Link
href="http://documentation.sas.com/?softwareId=STUDIOMID&amp;softwareVersion=3.8&amp;softwareContextId=tasks&amp;requestor=inapp">SAS Studio Task Reference Guide</Link>
```

```
        </Links>
        </Registration>

        <Metadata>

                <DataSources>
                </DataSources>

                <Options>
                </Options>

        </Metadata>

        <UI>
        </UI>

        <CodeTemplate>
                <![CDATA[

proc print data=sashelp.cars;run;

                ]]>
        </CodeTemplate>
</Task>
```

Table 3. NEW TASK template generated by SAS Studio to get started with the custom task.

WATERFALL PLOT

The following task adopted a "NEW TASK" template displayed above in *Table 3* and modified it to create a waterfall plot. The "DATA" tab shown in *Figure 6* inputs the dataset and required variables to generate the waterfall plot. As described in *Figure 7*, the "APPEARANCE" tab includes options to change the title, subtitle, color scheme, legend location, position and orientation, and values on Y-Axis (maximum, minimum, and increment).

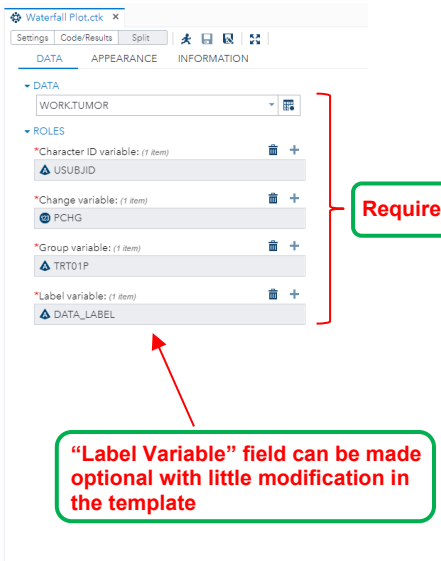

Figure 6. Input dataset and variables specified in the "DATA" tab to generate the code.

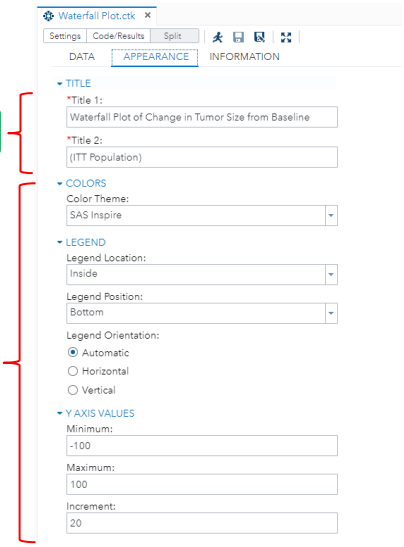

Figure 7. Modified title and subtitle as required in the "APPEARANCE" tab.

The SAS studio generates the code simultaneously when the user provides the input. The following SAS code is generated after entering all the required elements and options. Now users can click on the “Run” icon or press “F3” to run the code and generate the waterfall plot, as displayed in *Figure 9*.

CODE LOG RESULTS

```
/*
 *
 * Task code generated by SAS Studio 3.8
 *
 *
 */

%let fillcolors= cx21b9b7 cx4141e0 cx7db71a cx8e2f8a cxd38506 cx0abf85 cx2f90ec;
title 'Waterfall Plot of Change in Tumor Size from Baseline';
title2 '(ITT Population)';

proc sgplot data=WORK.TUMOR nowall noborder;
    styleattrs datacolors=(&fillcolors) datacontrastcolors=(black);
    vbar USUBJID / response=PCHG group=TRT01P categoryorder=respdesc
        datalabel=DATA_LABEL datalabelattrs=(size=5 weight=bold) groupdisplay=cluster
        clusterwidth=1;
    refline 20 -30/ lineattrs=(pattern=shortdash);
    xaxis display=none;
    yaxis values=(-100 to 100 by 20);
    keylegend / title='' location=inside position=bottom border;
run;

```

Figure 8. SAS Code Generated for waterfall plot with the custom task.

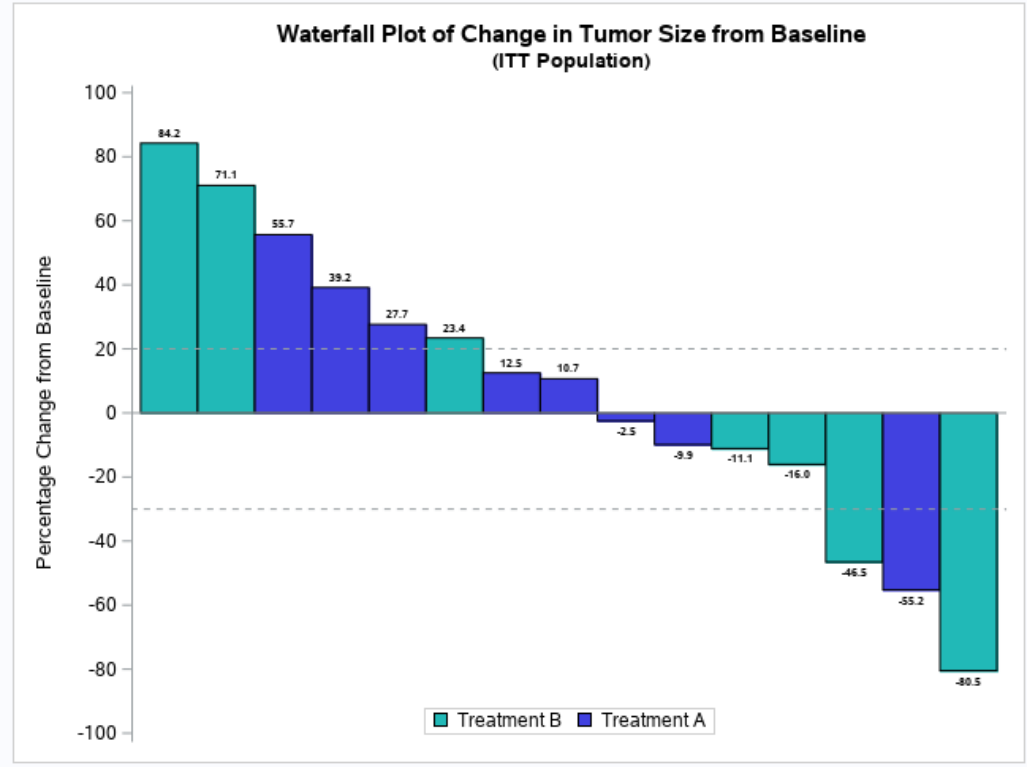

Figure 9. Waterfall plot with data label.

CONCLUSION

SAS Studio offers a lot of built-in tasks and provides flexibility to create a custom task as per the user requirements. It helps new and experienced programmers to create a SAS code with ease and provide flexibility. The approach reduces the programming efforts and time, especially dealing with complex programming, and generates high-quality outputs. Additionally, working with the task can also increase work efficiency and generate repeated outputs with different subgroup analysis in a few clicks. This paper provides an overview of task basics, standard point-and-click tasks, and creating a custom task by adopting and modifying a standard or new template.

REFERENCES AND RECOMMENDED READING

- SAS Institute. "SAS® Studio: User's Guide" Retrieved from https://documentation.sas.com/doc/en/webeditorcdc/v_009/webeditorug/titlepage.htm
- SAS Institute. "Overview of SAS Studio Tasks" Retrieved from https://documentation.sas.com/doc/en/webeditorcdc/5.2/webeditortaskgs/n0zdfvlmgqlrsun1gy7frid4oa0h.htm
- SAS Institute. "Built-in Tasks and Associated SAS Procedures and Licenses" Retrieved from https://documentation.sas.com/doc/en/egdoccdc/8.1/egamotasks/n0dgacov10xz69n1dttxmq4qntbj.htm
- SAS Institute. "SAS Studio 5.2: Developer's Guide to Writing Custom Tasks" Retrieved from https://go.documentation.sas.com/api/docsets/webeditordg/5.2/content/webeditordg.pdf
- SAS Institute. "Developing Your Own SAS Studio Custom Tasks for Advanced Analytics" Retrieved from https://support.sas.com/resources/papers/proceedings17/SAS0677-2017.pdf

ACKNOWLEDGMENTS

The authors would like to acknowledge colleagues who provided their feedback and helped us finalize this paper. In particular, we would like to thank Jeff Xia and Amy Gillespie for proofreading and providing critical comments.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Himanshu Patel
Enterprise: Merck & Co., Inc.
Address: 126 E. Lincoln Ave
City, State & Zip: Rahway, NJ 07065-4607
Phone: (732) 594-4071
E-mail: himanshu.patel@merck.com
Web: www.merck.com

Name: Chintan Pandya
Enterprise: Merck & Co., Inc.
Address: 126 E. Lincoln Ave
City, State & Zip: Rahway, NJ 07065-4607
Phone: (732) 594-7117
E-mail: chintan.pandya@merck.com
Web: www.merck.com