<table>
<thead>
<tr>
<th>TSC Category</th>
<th>Big Data Analytics</th>
</tr>
</thead>
</table>

**TSC**

Data Analytics System Design

**TSC Description**

Integrate the use of data analytics in the production environment for the identification of bottlenecks and system improvements

<table>
<thead>
<tr>
<th>TSC Proficiency Description</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRE-DAT-3005-1.1</td>
<td>PRE-DAT-4005-1.1</td>
<td>PRE-DAT-5005-1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingest and prepare the data for big data analytics by reviewing the data requirements and cleansing the data required for the analytics</td>
<td>Review the requirements of the statistical model to ensure it stays aligned with business needs and deploy the model to the production environment for user’s operational use</td>
<td>Define the hypotheses for the business problem, select the big data technologies and tools to be implemented in an organisation based on the data requirements</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Knowledge**

- Tools and/or programming languages for ingesting and/or transforming and/or cleansing big data
- Nature of data and data sources of the data to be prepared
- Organisation’s data collection process
- Concepts of data quality
- Data modelling
- Considerations of the analytics architecture
- Analytics architecture
- Analytical tools or data warehouse
- Tools and techniques for hypothesis formulation
- Components of different big data technologies and tools
- Pros and cons of different big data technologies and tools
- Types of big data frameworks
- Data requirements required for analytics
- Data analytics plan

**Abilities**

- Review the data requirements required for the analytics project
- Ingest data from different data sources into the analytics platform using the tools and/or programming language
- Cleanse and transform the data according to the data requirements to support the analytics project
- Resolve and follow up with any issues arising
- Select the runtime environment for the statistical model to be deployed and user requirements with the relevant stakeholders
- Define the analytics architecture requirements with the IT team to deploy the statistical model
- Develop the process to support the operations of the model with relevant stakeholders
- Monitor and tune the deployed model to ensure
- Define the business problem with the business stakeholders
- Formulate the hypotheses based on the business problem
- Evaluate and select the appropriate Big Data technologies and tools
- Design and drive the solution based on the business problem and hypotheses
| | during the data preparation | that it delivers the expected outcomes and aligns with the business changes | |