

Cheat Sheet for comprehensive Oracle Certified Professional, Big Data

Data Ingestion and Integration

Oracle Data Integrator (ODI)

- Key Features:

- **ETL/ELT:** Supports both ETL (Extract, Transform, Load) and ELT (Extract, Load, Transform) processes.

- **Data Quality:** Built-in data quality checks and transformations.

- **Scalability:** Handles large volumes of data with high performance.

- Shortcuts:

- **Creating a New Mapping:** `Ctrl + N`

- **Running a Scenario:** `F9`

- **Debugging:** `Ctrl + Shift + D`

- Tips:

- Use **Knowledge Modules** for reusable transformations.

- Schedule jobs using **OdiSchedule** for automated data loads.

Oracle GoldenGate

- Key Features:

- **Real-Time Data Replication:** Supports real-time data replication across heterogeneous systems.

- **Change Data Capture (CDC):** Captures and replicates changes in real-time.

- **High Availability:** Ensures data availability and consistency.

- Shortcuts:

- **Start Capture Process:** `ggsci> start mgr`

- **Add Trail File:** `ggsci> add trandata <schema>.<table>`

- Tips:

- Use **Extract** and **Replicat** processes for data movement.
- Configure **Checkpoints** to manage data consistency.

Data Storage and Management

Oracle Big Data SQL

- Key Features:

- **Unified Query:** Query data across Oracle Database, Hadoop, and NoSQL.
- **Performance:** Optimized for querying large datasets.
- **Security:** Integrated with Oracle Database security features.

- Shortcuts:

- **Query Execution:** `Ctrl + Enter`
- **Explain Plan:** `EXPLAIN PLAN FOR <query>`

- Tips:

- Use **External Tables** for querying Hadoop data.
- Leverage **Oracle Advanced Analytics** for data mining.

Oracle NoSQL Database

- Key Features:

- **Scalability:** Horizontally scalable for large datasets.
- **Consistency:** Supports strong consistency and eventual consistency.
- **High Availability:** Built-in replication and failover mechanisms.

- Shortcuts:

- **Create Table:** `CREATE TABLE <table_name> (key <key_type>, column <column_type>, ...)`
- **Insert Data:** `INSERT INTO <table_name> (key, column) VALUES (value1, value2)`

- Tips:

- Use **KVStore** for distributed data storage.
- Configure **Replication** for high availability.

Data Processing and Analytics

Oracle Big Data Connectors

- Key Features:

- **HDFS Connector:** Access and manage data in Hadoop Distributed File System (HDFS).
- **Spark Connector:** Integrate with Apache Spark for big data processing.
- **Kafka Connector:** Consume and produce data from Apache Kafka.

- Shortcuts:

- **HDFS Access:** ``hadoop fs -ls /path/to/directory``
- **Spark Job:** ``spark-submit --class <class_name> <jar_file>``

- Tips:

- Use **Oracle Big Data Lite VM** for development and testing.
- Integrate with **Oracle Cloud Infrastructure** for scalable processing.

Oracle Big Data Appliance

- Key Features:

- **Integrated Solution:** Pre-configured hardware and software for big data workloads.
- **Performance:** Optimized for high-performance data processing.
- **Management:** Simplified management and monitoring.

- Shortcuts:

- **Cluster Status:** ``nodetool status``
- **Resource Management:** ``yarn application -list``

- Tips:

- Use **Oracle Big Data Spatial and Graph** for geospatial and graph analytics.
- Monitor cluster health using **Oracle Enterprise Manager**.

Data Visualization and Reporting

Oracle Business Intelligence (BI)

- Key Features:

- **Dashboards:** Create interactive dashboards for data visualization.
- **Reports:** Generate detailed reports with customizable templates.
- **Mobile Access:** Access BI content on mobile devices.
- **Shortcuts:**
 - **Create Dashboard:** `Ctrl + D`
 - **Export Report:** `Ctrl + E`
- **Tips:**
 - Use **Oracle BI Publisher** for advanced reporting.
 - Integrate with **Oracle Analytics Cloud** for cloud-based BI.

Oracle Data Visualization

- **Key Features:**
 - **Interactive Visualizations:** Create interactive charts, graphs, and maps.
 - **Data Discovery:** Explore data through guided and ad-hoc analysis.
 - **Collaboration:** Share visualizations and insights with team members.
- **Shortcuts:**
 - **Create Visualization:** `Ctrl + V`
 - **Filter Data:** `Ctrl + F`
- **Tips:**
 - Use **Oracle Data Visualization Desktop** for offline analysis.
 - Leverage **Oracle Machine Learning** for predictive analytics.

Security and Compliance

Oracle Data Masking and Subsetting

- **Key Features:**
 - **Data Masking:** Mask sensitive data to protect privacy.
 - **Data Subsetting:** Create subsets of data for development and testing.
 - **Compliance:** Ensure compliance with data protection regulations.

- **Shortcuts:**

- **Create Masking Set:** `CREATE MASKING SET <set_name>`
- **Apply Masking:** `APPLY MASKING SET <set_name>`

- **Tips:**

- Use **Oracle Data Safe** for comprehensive data security.
- Regularly update masking rules to reflect changing data.

Oracle Audit Vault and Database Firewall

- **Key Features:**

- **Audit Collection:** Collect and analyze audit data from multiple sources.
- **Database Firewall:** Protect databases from unauthorized access and SQL injection.
- **Compliance Reporting:** Generate reports for regulatory compliance.

- **Shortcuts:**

- **Start Audit Collection:** `avcli start collection`
- **Configure Firewall Rules:** `avcli configure firewall`

- **Tips:**

- Use **Oracle Audit Vault** for centralized audit management.
- Regularly review firewall rules to ensure security.

Performance Tuning and Optimization

Oracle Big Data SQL Tuning

- **Key Features:**

- **Query Optimization:** Optimize queries for performance.
- **Indexing:** Use indexes to speed up query execution.
- **Caching:** Leverage caching mechanisms for frequently accessed data.

- **Shortcuts:**

- **Explain Plan:** `EXPLAIN PLAN FOR <query>`
- **Index Creation:** `CREATE INDEX <index_name> ON <table_name> (<column_name>)`

- **Tips:**

- Use **Oracle SQL Developer** for query tuning.
- Monitor performance using **Oracle Enterprise Manager**.

Oracle Big Data Appliance Tuning

- **Key Features:**

- **Resource Management:** Optimize resource allocation for big data workloads.
- **Cluster Tuning:** Tune cluster settings for performance.
- **Monitoring:** Monitor cluster performance and health.

- **Shortcuts:**

- **Resource Management:** ``yarn application -list``
- **Cluster Status:** ``nodetool status``

- **Tips:**

- Use **Oracle Big Data Lite VM** for development and testing.
- Regularly review and optimize cluster configurations.

Cloud Integration

Oracle Cloud Infrastructure (OCI)

- **Key Features:**

- **Scalability:** Scale resources up or down based on demand.
- **Security:** Built-in security features and compliance certifications.
- **Integration:** Seamless integration with on-premises and other cloud environments.

- **Shortcuts:**

- **Create Instance:** ``oci compute instance launch``
- **Monitor Resources:** ``oci monitoring metric-data summarize-metrics-data``

- **Tips:**

- Use **Oracle Cloud Infrastructure Data Integration** for ETL in the cloud.
- Leverage **Oracle Cloud Infrastructure Data Flow** for Apache Spark jobs.

Oracle Autonomous Database

- **Key Features:**

- **Autonomous:** Self-managing, self-securing, and self-repairing.
- **Performance:** Optimized for high performance and scalability.
- **Security:** Built-in security features and compliance certifications.

- **Shortcuts:**

- **Create Database:** ``oci db autonomous-database create``
- **Monitor Performance:** ``oci db autonomous-database get``

- **Tips:**

- Use **Oracle Autonomous Data Warehouse** for data warehousing.
- Integrate with **Oracle Analytics Cloud** for advanced analytics.

Additional Resources

Documentation and Support

- **Oracle Documentation:** [Oracle Documentation](<https://docs.oracle.com/>)
- **Oracle Support:** [Oracle Support](<https://support.oracle.com/>)
- **Oracle Community:** [Oracle Community](<https://community.oracle.com/>)

Training and Certification

- **Oracle University:** [Oracle University](<https://education.oracle.com/>)
- **Certification Path:** [Oracle Certification](<https://education.oracle.com/certification>)

Tools and Utilities

- **Oracle SQL Developer:** [Oracle SQL Developer](<https://www.oracle.com/tools/downloads/sqldev-downloads.html>)
- **Oracle Enterprise Manager:** [Oracle Enterprise Manager](<https://www.oracle.com/enterprise-manager/>)
- **Oracle Big Data Lite VM:** [Oracle Big Data Lite VM](<https://www.oracle.com/technetwork/database/bigdata-appliance/oracle-bigdatalite-2104726.html>)

Conclusion

This cheat sheet provides a comprehensive overview of key features, shortcuts, tips, and tricks for Oracle Certified Professional, Big Data. Use these resources to enhance your skills and effectively manage big data environments.

By Ahmed Baheeg Khorshid

ver 1.0