

Data Visualization - Exam Questions

Objective Questions (MCQ/MSQ) - 20 Questions

1. Which Seaborn function creates a correlation heatmap?

- A) `sns.heatmap()`
- B) `sns.corrplot()`
- C) `sns.correlation()`
- D) `sns.matrix()`

Answer: A

2. In RevealJS, which HTML attribute enables Markdown content?

- A) `data-content="markdown"`
- B) `data-markdown`
- C) `markdown="true"`
- D) `content-type="markdown"`

Answer: B

3. What is the primary advantage of Marimo over Jupyter notebooks?

- A) Better performance
- B) Reactive execution and dependency tracking
- C) More visualization options
- D) Easier installation

Answer: B

4. Which Excel feature creates in-cell trend visualizations?

- A) Charts
- B) Sparklines
- C) Conditional formatting
- D) Data bars

Answer: B

5. In PowerPoint, which transition enables smooth animations between slides?

- A) Fade
- B) Morph
- C) Slide
- D) Push

Answer: B

6. Which RAWGraphs chart type is best for showing hierarchical data?

- A) Bar chart
- B) Scatter plot
- C) Treemap
- D) Line chart

Answer: C

7. What does “object constancy” mean in animated visualizations?

- A) Objects maintain the same color
- B) Objects maintain the same size
- C) Objects maintain their identity across frames
- D) Objects move at constant speed

Answer: C

8. Which Flourish feature controls the timing of element appearances in animations?

- A) Duration
- B) Transition
- C) Stagger
- D) Delay

Answer: C

9. In Kumu, what format is required for network data upload?

- A) Adjacency matrix
- B) From-node to-node with connection strength
- C) JSON format
- D) CSV with headers

Answer: B

10. Which Seaborn function creates small multiple plots?

- A) `sns.subplot()`
- B) `sns.FacetGrid()`
- C) `sns.multiple()`
- D) `sns.grid()`

Answer: B

11. What is the recommended approach for handling overplotting in scatter plots?

- A) Use smaller points
- B) Use alpha transparency
- C) Use hexbin plots
- D) All of the above

Answer: D

12. Which color palette type should be used for categorical data?

- A) Sequential
- B) Diverging
- C) Qualitative
- D) Continuous

Answer: C

13. In Excel sparklines, which type shows win/loss patterns?

- A) Line
- B) Column
- C) Win/Loss
- D) Area

Answer: C

14. What does the `GROWTH()` function in Excel predict?

- A) Linear trends
- B) Exponential trends
- C) Seasonal patterns
- D) Random values

Answer: B

15. Which RevealJS plugin enables syntax highlighting for code?

- A) RevealCode
- B) RevealHighlight
- C) RevealSyntax
- D) RevealPrism

Answer: B

16. In Marimo, what happens when you change a cell that other cells depend on?

- A) Nothing
- B) You must manually run dependent cells
- C) Dependent cells update automatically
- D) The notebook restarts

Answer: C

17. Which ImageMagick command creates a montage of multiple images?

- A) `convert`
- B) `montage`
- C) `composite`
- D) `mogrify`

Answer: B

18. What is the primary purpose of data storytelling?

- A) Show all available data
- B) Make insights memorable and actionable
- C) Create complex visualizations
- D) Demonstrate technical skills

Answer: B

19. Which network analysis metric identifies nodes that bridge different communities?

- A) Degree centrality
- B) Closeness centrality
- C) Betweenness centrality
- D) Eigenvector centrality

Answer: C

20. In Flourish, which template allows morphing between different chart types?

- A) Bar Chart Race
- B) Line Chart Race
- C) Line Bar Pie
- D) Scatter Plot

Answer: C

Subjective/Scenario Questions - 20 Questions

1. Visualization Strategy Design

You need to present quarterly sales data to three different audiences: executives (high-level trends), sales managers (detailed performance), and analysts (statistical insights). Design appropriate visualizations for each audience and justify your choices.

Answer: Executives: Dashboard with KPI cards, trend lines, and geographic heat maps focusing on key metrics and year-over-year comparisons. **Sales Managers:** Interactive drill-down charts, performance rankings, and target vs. actual comparisons with filters by region/product. **Analysts:** Correlation matrices, statistical distributions, regression plots, and detailed data tables with export capabilities. Use consistent color schemes and progressive disclosure of detail.

2. Tool Selection for Interactive Presentation

Compare RevealJS, PowerPoint, and Flourish for creating an interactive data presentation about climate change trends. Consider audience engagement, technical requirements, and maintenance needs.

Answer: RevealJS: Best for technical audiences, supports live data integration, requires web development skills, highly customizable. **PowerPoint:** Familiar

interface, good for business audiences, limited interactivity, easy to maintain.

Flourish: Excellent for data stories, template-based, good balance of ease and interactivity. Recommend RevealJS for technical conferences, PowerPoint for board meetings, Flourish for public communication.

3. Animated Visualization Design

Create an animated visualization showing the evolution of COVID-19 cases across countries over time. Compare PowerPoint Morph vs. Flourish approaches and recommend the best strategy.

Answer: PowerPoint Morph: Manual control, custom animations, voiceover integration, time-intensive setup. **Flourish:** Automated race charts, professional templates, easy data updates, limited customization. Recommend Flourish for rapid prototyping and social media, PowerPoint for detailed storytelling with custom narrative. Consider hybrid approach: Flourish for initial concept, PowerPoint for final polished version.

4. Dashboard Architecture

Design a real-time business intelligence dashboard that updates automatically and serves multiple departments. What visualization components and technical architecture would you implement?

Answer: Architecture: (1) **Data layer** - streaming ETL pipeline, (2) **API layer** - real-time data endpoints, (3) **Visualization layer** - web-based dashboard with WebSocket updates. Components: KPI cards, trend charts, geographic maps, alert panels. Use **D3.js** or **Observable Plot** for custom visualizations, **React/Vue** for interactivity, **WebSocket** for real-time updates. Implement caching and progressive loading for performance.

5. Color Palette Strategy

Design a comprehensive color strategy for a data visualization system that must work for colorblind users and in both digital and print formats. What guidelines would you establish?

Answer: (1) **Accessibility:** Use colorblind-safe palettes (viridis, ColorBrewer), ensure sufficient contrast ratios, provide alternative encodings (patterns, shapes). (2) **Consistency:** Establish semantic color meanings, document color codes, create style guides. (3) **Adaptability:** Test in grayscale, optimize for different media, provide high-contrast alternatives. (4) **Tools:** Use ColorBrewer, Viz Palette, and accessibility checkers.

6. Network Visualization Optimization

You have a social network with 10,000 nodes and 50,000 edges that creates a “hairball” visualization. Design strategies to make it interpretable and actionable.

Answer: (1) **Filtering:** Show only high-degree nodes, filter by edge weights, implement search functionality. (2) **Layout:** Use force-directed algorithms, apply clustering, create hierarchical views. (3) **Interaction:** Enable zoom/pan, highlight neighborhoods on hover, provide detail-on-demand. (4) **Aggregation:** Group similar nodes, show community structures, use edge bundling. (5) **Multiple views:** Overview + detail, adjacency matrix alternative.

7. Performance Optimization

Your Seaborn visualizations take 30 seconds to render with a 1M row dataset. Design optimization strategies while maintaining visual quality.

Answer: (1) **Data sampling:** Statistical sampling for exploration, full data for final charts. (2) **Efficient plotting:** Use `matplotlib` directly for simple plots, leverage `datashader` for large datasets. (3) **Caching:** Cache processed data, use incremental updates. (4) **Optimization:** Optimize data types, use vectorized operations, consider `polars` for data processing. (5) **Progressive rendering:** Show previews while processing, implement lazy loading.

8. Cross-Platform Compatibility

Design a visualization system that works consistently across web browsers, mobile devices, and print media. What technical and design considerations would you address?

Answer: (1) **Responsive design:** Flexible layouts, scalable fonts, adaptive chart sizes. (2) **Technology stack:** SVG for scalability, progressive enhancement, fallback options. (3) **Testing:** Cross-browser testing, device testing, print preview validation. (4) **Performance:** Optimize for mobile networks, implement lazy loading, minimize dependencies. (5) **Accessibility:** Screen reader compatibility, keyboard navigation, high contrast modes.

9. Narrative Structure

Create a data storytelling framework for presenting complex analytical findings to non-technical stakeholders. What narrative techniques and visualization strategies would you use?

Answer: (1) **Structure:** Executive summary → context → findings → implications → recommendations. (2) **Techniques:** Start with familiar concepts, use analogies, progressive disclosure of complexity. (3) **Visuals:** Clear titles, annotations, consistent styling, minimal cognitive load. (4) **Engagement:** Interactive elements, questions, scenarios. (5) **Validation:** User testing, feedback loops, iterative refinement.

10. Real-time Monitoring Dashboard

Design a monitoring dashboard for a data pipeline that processes 1TB daily. What visualizations and alerting mechanisms would you implement?

Answer: (1) **Key metrics:** Throughput rates, error rates, latency percentiles, resource utilization. (2) **Visualizations:** Time series charts, status indicators, geographic distribution, queue depths. (3) **Alerting:** Threshold-based alerts, anomaly detection, escalation procedures. (4) **Interactivity:** Drill-down capabilities, filtering, historical comparisons. (5) **Performance:** Efficient data aggregation, caching strategies, progressive loading.

11. Accessibility Compliance

Ensure your data visualizations comply with WCAG 2.1 AA standards. What specific measures would you implement for different chart types?

Answer: (1) **Color:** Sufficient contrast ratios, alternative encodings beyond color. (2) **Text:** Scalable fonts, descriptive titles, alt text for images. (3) **Interaction:** Keyboard navigation, screen reader compatibility, focus indicators. (4) **Structure:** Logical reading order, semantic markup, data tables for complex charts. (5) **Testing:** Automated accessibility testing, user testing with assistive technologies.

12. Version Control for Visualizations

Implement version control for a team creating data visualizations. How would you handle code, data, and design assets while ensuring reproducibility?

Answer: (1) **Code:** Git for scripts and notebooks, semantic versioning, code reviews. (2) **Data:** DVC for dataset versioning, data lineage tracking, checksums for integrity. (3) **Assets:** Git LFS for images, style guides in version control, design system documentation. (4) **Reproducibility:** Environment management (Docker/conda), automated testing, CI/CD pipelines. (5) **Collaboration:** Branch strategies, merge procedures, documentation standards.

13. Mobile-First Visualization

Design a mobile-first approach for data visualizations that will primarily be viewed on smartphones. What design principles and technical considerations would you apply?

Answer: (1) **Design:** Simplified charts, vertical layouts, touch-friendly interactions, minimal text. (2) **Performance:** Optimized loading, progressive enhancement, offline capabilities. (3) **Interaction:** Swipe gestures, tap targets, voice navigation options. (4) **Responsive:** Breakpoint strategies, flexible grids, scalable typography. (5) **Testing:** Device testing, performance monitoring, user experience validation.

14. Automated Insight Generation

Create a system that automatically generates narrative insights from data visualizations. What natural language generation and analysis techniques would you implement?

Answer: (1) **Pattern detection:** Statistical anomalies, trend identification, correlation discovery. (2) **NLG techniques:** Template-based generation, neural language models, context-aware descriptions. (3) **Prioritization:** Significance scoring, business impact assessment, user preference learning. (4) **Validation:** Fact-checking, confidence intervals, human review processes. (5) **Personalization:** User role adaptation, customizable insights, feedback incorporation.

15. Collaborative Visualization Platform

Design a platform where multiple stakeholders can collaboratively create and review data visualizations. What features and workflows would you implement?

Answer: (1) **Collaboration:** Real-time editing, comment systems, version history, role-based permissions. (2) **Workflow:** Review processes, approval workflows, publication pipelines. (3) **Integration:** Data source connections, export capabilities, embedding options. (4) **Governance:** Style guides, template libraries, quality standards. (5) **Communication:** Notification systems, discussion threads, change tracking.

16. Performance Benchmarking

Establish performance benchmarks for different visualization libraries and chart types. What metrics would you track and how would you conduct fair comparisons?

Answer: (1) **Metrics:** Rendering time, memory usage, file size, interaction responsiveness. (2) **Test scenarios:** Various data sizes, chart complexities, device capabilities. (3) **Methodology:** Controlled environments, statistical significance, multiple runs. (4) **Libraries:** Compare D3.js, Chart.js, Plotly, Observable Plot, native canvas. (5) **Reporting:** Performance dashboards, regression testing, optimization recommendations.

17. Ethical Visualization Guidelines

Develop ethical guidelines for data visualization to prevent misleading representations. What principles and review processes would you establish?

Answer: (1) **Principles:** Truthful representation, appropriate scales, clear context, uncertainty communication. (2) **Guidelines:** Avoid truncated axes, consistent scales, representative sampling, bias disclosure. (3) **Review process:** Peer review, stakeholder validation, ethical checklists. (4) **Training:** Team education, best practices documentation, case study analysis. (5) **Monitoring:** Feedback collection, impact assessment, continuous improvement.

18. Internationalization Strategy

Design a visualization system that works across different languages, cultures, and data formats. What localization considerations would you address?

Answer: (1) **Language:** Text translation, right-to-left layouts, font selection, character encoding. (2) **Cultural:** Color meanings, reading patterns, cultural sensitivities, local conventions. (3) **Data formats:** Date/time formats, number formatting, currency symbols, measurement units. (4) **Technical:** Unicode support, font loading, layout flexibility. (5) **Testing:** Native speaker review, cultural validation, usability testing.

19. Disaster Recovery for Visualization Systems

Design a disaster recovery plan for a critical business intelligence dashboard system. What backup and recovery strategies would you implement?

Answer: (1) **Infrastructure:** Multi-region deployment, automated backups, failover mechanisms. (2) **Data:** Regular snapshots, incremental backups, data validation procedures. (3) **Code:** Version control, deployment automation, rollback procedures. (4) **Monitoring:** Health checks, alert systems, recovery time tracking. (5) **Testing:** Regular drills, recovery validation, documentation updates.

20. ROI Measurement for Visualization Projects

Develop a framework to measure and demonstrate the return on investment for data visualization initiatives. What metrics and methodologies would you use?

Answer: (1) **Quantitative metrics:** Decision speed improvement, error reduction, time savings, cost avoidance. (2) **Qualitative measures:** User satisfaction, adoption

rates, stakeholder feedback. (3) **Business impact:** Revenue attribution, process improvements, strategic insights. (4) **Measurement methods:** A/B testing, before/after comparisons, user analytics. (5) **Reporting:** Executive dashboards, case studies, success stories, continuous monitoring.