

工业和信息化人才培养工程大数据分析师

复习题库（一）

- 1、当前大数据技术的基础是由（C）首先提出的。
A: 微软 B: 百度 C: 谷歌 D: 阿里巴巴
- 2、大数据的起源是（C）
A: 金融 B: 电信 C: 互联网 D: 公共管理
- 3、根据不同的业务需求来建立数据模型，抽取最有意义的向量，决定选取哪种方法的数据分析角色人员是（C）。
A: 数据管理人员 B: 数据分析员 C: 研究科学家 D: 软件开发工程师
- 4、（D）反映数据的精细化程度，越细化的数据，价值越高。
A: 规模 B: 活性 C: 关联度 D: 颗粒度
- 5、数据清洗的方法不包括（D）。
A: 缺失值处理 B: 噪声数据清除 C: 一致性检查 D: 重复数据记录处理
- 6、智能健康手环的应用开发，体现了（D）的数据采集技术的应用。
A: 统计报表 B: 网络爬虫 C: API接口 D: 传感器
- 7、下列关于数据重组的说法中，错误的是（A）。
A: 数据重组是数据的重新生产和重新采集
B: 数据重组能够使数据焕发新的光芒
C: 数据重组实现的关键在于多源数据融合和数据集成
D: 数据重组有利于实现新颖的数据模式创新
- 8、智慧城市的构建，不包含（C）。
A: 数字城市 B: 物联网 C: 联网监控 D: 云计算
- 9、大数据的最显著特征是（A）。
A: 数据规模大 B: 数据类型多样 C: 数据处理速度快 D: 数据价值密度高
- 10、美国海军军官莫里通过对前人航海日志的分析，绘制了新的航海路线图，标明了大风与洋流可能发生的地点。这体现了大数据分析理念中的（B）。
A: 在数据基础上倾向于全体数据而不是抽样数据
B: 在分析方法上更注重相关分析而不是因果分析
C: 在分析效果上更追究效率而不是绝对精确
D: 在数据规模上强调相对数据而不是绝对数据
- 11、下列关于舍恩伯格对大数据特点的说法中，错误的是（D）。
A: 数据规模大 B: 数据类型多样 C: 数据处理速度快 D: 数据价值密度高
- 12、当前社会中，最为突出的大数据环境是（A）。
A: 互联网 B: 物联网 C: 综合国力 D: 自然资源
- 13、在数据生命周期管理实践中，（B）是
A: 数据存储和备份规范 B: 数据管理和维护

- C: 数据价值发觉和利用 D: 数据应用开发和管理
- 14、 下列关于网络用户行为的说法中， 错误的是（ C ）。
- A: 网络公司能够捕捉到用户在其网站上的所有行为
B: 用户离散的交互痕迹能够为企业提升服务质量提供参考
C: 数字轨迹用完即自动删除
D: 用户的隐私安全很难得以规范保护
- 15、 下列关于计算机存储容量单位的说法中， 错误的是（ C ）。
- A: $1KB < 1MB < 1GB$ B: 基本单位是字节（Byte）
C: 一个汉字需要一个字节的存储空间 D: 一个字节能够容纳一个英文字符，
- 16、 下列关于聚类挖掘技术的说法中， 错误的是（ B ）。
- A: 不预先设定数据归类类目， 完全根据数据本身性质将数据聚合成不同类别
B: 要求同类数据的内容相似度尽可能小
C: 要求不同类数据的内容相似度尽可能小
- 17、 下列国家的大数据发展行动中， 集中体现“重视基础、 首都先行”的国家是（ D ）。
- A: 美国 B: 日本 C: 中国 D: 韩国
- 18、 下列关于大数据的分析理念的说法中， 错误的是（ D ）。
- A: 在数据基础上倾向于全体数据而不是抽样数据
B: 在分析方法上更注重相关分析而不是因果分析
C: 在分析效果上更追究效率而不是绝对精确
D: 在数据规模上强调相对数据而不是绝对数据
- 19、 万维网之父是（ C ）。
- A: 彼得·德鲁克 B: 舍恩伯格 C: 蒂姆·伯纳斯-李 D: 斯科特·布朗
- 20、 Mac OS 系统的开发者是（ C ）。
- A: 微软公司 B: 惠普公司 C: 苹果公司 D: IBM 公司
- 21、 大数据时代， 数据使用的关键是（ D ）。
- A: 数据收集 B: 数据存储 C: 数据分析 D: 数据再利用
- 22、 下列关于数据交易市场的说法中， 错误的是（ C ）。
- A: 数据交易市场是大数据产业发展到一定程度的产物
B: 商业化的数据交易活动催生了多方参与的第三方数据交易市场
C: 数据交易市场通过生产数据、 研发和分析数据， 为数据交易提供帮助
D: 数据交易市场是大数据资源化的必然产物
- 23、 下列论据中， 能够支撑“大数据无所不能”的观点的是（ A ）。
- A: 互联网金融打破了传统的观念和和行为 B: 大数据存在泡沫
C: 大数据具有非常高的成本 D: 个人隐私泄露与信息安全担忧
- 24、 数据仓库的最终目的是（ D ）。
- A: 收集业务需求 B: 建立数据仓库逻辑模型
C: 开发数据仓库的应用分析 D: 为用户和业务部门提供决策支持
- 25、 支撑大数据业务的基础是（ B ）。

A: 数据科学 B: 数据应用 C: 数据硬件 D: 数据人才

26、以下选项中，不属于信息时代的定律的是（ ）。

- A. 吉尔德定律 B. 摩尔定律 C. 麦特卡尔夫定律 D. 达律多定律

27. 根据周琦老师所讲，高德交通信息服务覆盖全国高速（ ）以上。

- A. 60% B. 90% C. 70% D. 50%

28. 根据周琦老师所讲，以下哪项不属于数据挖掘的内容？

- A. 补充与完善路网属性 B. 多维分析统计用户出行规律
 C. 高德地图导航有躲避拥堵功能 D. 建立道路拥堵概率与拥堵趋势变化模型

29. 2012 年全国各城市支付宝人均支出排名中，位居第二位的是（ ）

- A. 杭州市 B. 嘉兴市 C. 高雄市 D. 嘉义市

30. 第一个提出大数据概念的公司是（ ）。

- A. 微软公司 B. 谷歌公司 C. 脸谱公司 D. 麦肯锡公司

31.“（ ）阿里巴巴·贵州年货节”销售额突破 8.5 亿元，促进了贵州电子商务加快发展。

- A. 2015 B. 2016 C. 2013 D. 2014

32. 关于大数据在社会综合治理中的作用，以下理解不正确的是（ ）。

- A. 大数据的运用能够维护社会治安 B. 大数据的运用能够加强交通管理
 C. 大数据的运用能够杜绝抗生素的滥用 D. 大数据的运用有利于走群众路线

33. 根据周琦老师所讲，大数据加速道路网络快速更新，高德（ ）完成全国 10 万公里 15 万处更新。

- A. 2006 年 B. 2014 年 C. 2008 年 D. 2010 年

34. 第一个提出大数据概念的公司是

- A. 谷歌公司 B. 微软公司 C. 脸谱公司 D. 麦肯锡公司

35. 以下选项中，不属于大数据对人才能力的要求是

- A. 业务能力 B. 数学统计能力 C. IT 技术能力 D. 逻辑思维能力

36. 根据周琦老师所讲，高德交通报告针对全国（ ）个城市交通状态进行挖掘分析。

- A. 30 B. 21 C. 25 D. 38

37. 以下选项中，不属于大数据对人才能力的要求是（ ）。

- A. 业务能力 B. 数学统计能力 C. 逻辑思维能力 D. IT 技术能力

38.Web2.0 强调（）。

- A.机构 B.单位 C.个人 D.网站

39.根据周琦老师所讲，高德交通日均采集数十亿定位请求，系统处理月均（）公里驾驶里程覆盖。

- A.100 亿 B.150 亿 C.50 亿 D.80 亿

40.大数据正快速发展为对数量巨大、来源分散、格式多样的数据进行采集、存储和关联分析，从中发现新知识、创造新价值、提升新能力的（）。

- A.新一代技术平台 B.新一代信息技术和服务业态
- C.新一代服务业态 D.新一代信息技术