

必修一 数据与计算

专题一 数据与信息

【学习目标】

1. 理解数据的概念及表现形式。
2. 理解信息的概念及特征。
3. 理解知识的概念。
4. 了解数据、信息与知识的关系。
5. 理解大数据的概念及特征。
6. 了解大数据的应用与影响。
7. 熟悉进制数（二进制、八进制、十进制、十六进制）及不同进制数（整数）与十进制数的互相转换。
8. 掌握信息存储容量单位及其换算。
9. 了解编码的意义与作用。
10. 掌握ASCII码的基本知识。
11. 了解汉字编码（国标码）。
12. 理解模拟信号（模拟数据）及数字信号（数字数据）的概念。
13. 理解声音编码过程及声音存储容量的计算。
14. 理解图像编码及图像存储容量的计算。
15. 了解常见的文本、声音、图像、视频等文件格式。

第一课 数据、信息与知识

【知识梳理】

一、数据

1. 数据的概念

数据是现实世界客观事物的符号记录，是信息的载体；而信息是数据

的内涵（即含义），是对数据的语义解释。如“小李的身高是170厘米”这句话描述的是信息，其中“170”属于数据。

2. 数据的分类

（1）数据分为数值型数据与非数值型数据。数值型数据，如表示人的年龄、身高与体重的数据；非数值型数据，如文字、图像、声音和视频等数据。

（2）数据分为数字数据（二进制数）与模拟数据（如符号、语言、文字、图像、动画、声音和视频等）。例如，书上的文字、上课的铃声、漫画、电影等都属于模拟数据。

3. 数据的特点

有些数据是固定的，如电话区位号、邮政编码等；有些数据是会变化的，如气温、人的身高与体重、物品价格等；有些数据是随机的，如彩票开奖号码、抛骰子得到的号码等。

4. 学会找数据规律

有些数据是有规律的，如3，4，7，11的下一个数是18；2，5，8，11的下一个数是14。

5. 大数据

（1）大数据是指无法在可承受的时间范围内用常规软件工具进行高效捕捉、管理和处理的数据集合。大数据的产生、发展与互联网紧密相关。

（2）大数据特征：数据体量大（数据规模大）、数据类型多（数据多样且变化快）、处理速度快、应用价值高但价值密度低。

（3）大数据的应用：大数据已成为人们提取信息、作出决策的重要依据，如电商的个性化推荐、导航地图、行程码、掌上公交等都用到大数据。

（4）大数据的采集方法：网络采集法（网络爬虫）、传感器采集法、系统日志采集法等。

（5）大数据处理过程：数据采集与预处理→数据存储→数据挖掘→数据呈现。

二、信息

1. 信息的定义

信息的定义有三种：香农认为信息是“有新知识或内容的消息，信息

可以消除认识上的不确定性，信息可以使不确定量减少”；维纳认为“信息是一种资源”，信息、物质、能量是人类赖以生存的三大资源；钟义信认为“信息是事物运动的状态与状态变化的方式”。

信息是数据中所包含的意义，信息必须依附于某种载体，通过某种数据形式才能存储、表达和传播。典型的信息有信号、通知、资讯、消息、短信、广告、新闻、市场行情、天气预报等。

2. 信息的特征

信息常见的特征及其实例如下表所示。

信息特征	实例
普遍性	信息无处不在，无时不有。
依附性	学校通过 LED 屏幕显示的文字传达校园新闻。
可处理性	气象局对卫星收集到的气象数据进行分析，预测近期的天气。
传递性	北京冬奥会开幕式通过卫星电视向全球转播。
共享性	红绿灯显示的交通信号让每个路过的司机都能看到。
时效性	微信红包在发出后若 24 小时内未被领取，将自动退回。
价值性	我们去一个陌生的地方时，利用手机导航可以防止迷路。
真伪性	历史典故“空城计”和“减灶退敌”；手机诈骗短信。
价值相对性	酒驾扣 12 分这条信息对司机很重要，对小学生却用处不大。

3. 信息的载体

信息的载体主要有符号、语言、文字、图像、声音、动画、视频等。

三、知识

知识是人们运用大脑对获取或累积的信息进行系统提炼、研究和分析的结果，能够精确地反映事物的本质。知识是人们在改造世界的实践活动中获得的可用于指导实践的认识、规律和经验。知识是认识世界的结果，也是改造世界的依据。如感冒指数、身体健康指数，数理化中的公式和定理等属于知识；“标准气压下，冰水混合物的温度是0℃”描述的也是知识。

四、数据、信息、知识的关系

数据、信息、知识是人类认知客观世界由低到高的三个阶段，三者之间既有区别，又有密切联系。它们的关系可表示为数据→信息→知识。

【单元练习】

1. 数据是指()。
 - A. 数值的大小
 - B. 计量物体数量和质量的数值
 - C. 数字0, 1, 2, ..., 9及正、负号
 - D. 描述和记录客观事物属性的符号
2. 以下关于数据的说法, 不合理的是()。
 - A. 数据是信息的符号表示或载体
 - B. 数据分为数值数据和非数值数据
 - C. 只有计算机能识别的对象才是数据
 - D. 数据本身没有意义, 把数据加工成信息才体现其意义
3. 在日常生活中, 各地区的固定电话区号不会发生变化。这一事例说明有些数据是()。
 - A. 会消失的
 - B. 随机出现的
 - C. 固定不变的
 - D. 会变化的
4. 找数字规律: 1, 5, 6, 11, 17的下一个数字是()。
 - A. 23
 - B. 26
 - C. 22
 - D. 28
5. 计算机能处理的数据有()。
 - A. 只有二进制数
 - B. 只有数字和文本
 - C. 只有十进制数
 - D. 数字、文本、图像、声音、视频等
6. 以下关于大数据的说法, 错误的是()。
 - A. 大数据的处理离不开云计算
 - B. 从网络上搜索得到的数据称为大数据
 - C. 大数据在电商、电信、金融等行业都有巨大的社会价值和产业空间
 - D. 大数据是指无法在可承受的时间范围内用常规软件工具进行高效捕捉、管理和处理的数据集合
7. 大数据的特征包括()。
 - ①体量大 ②快速化 ③多样性 ④价值密度低
 - A. ①②③
 - B. ①③④
 - C. ②③④
 - D. ①②③④

8. 大数据起源于()。
- A. 电信网 B. 广电网 C. 企业网 D. 互联网
9. 以下应用大数据技术的是()。
- A. 健康码 B. 电子邮件
C. 文件传输 D. 微信私聊记录
10. 以下对信息的描述, 错误的是()。
- A. 信息可以消除认识上的不确定性
B. 信息是已经存在和被利用的消息
C. 信息是数据、消息中所包含的意义
D. 信息是事物的运动状态及其状态变化的方式
11. 以下不属于信息的是()。
- A. 书报刊 B. 会议通知 C. 天气预报 D. 股市行情
12. 以下属于信息的是()。
- A. 二维码 B. 计算机 C. 手机短信 D. 微视频
13. 以下不属于信息特征的是()。
- A. 依附性 B. 传递性 C. 共享性 D. 多元性
14. 蔬菜的价格会随着每天市场行情变化而波动, 这主要说明信息具有()。
- A. 不稳定性 B. 多样性 C. 时效性 D. 不确定性
15. 孙膑“减灶退敌”和诸葛孔明上演的“空城计”, 这两个历史典故主要说明了信息具有()。
- A. 传递性 B. 真伪性 C. 时效性 D. 局限性
16. 有些人通过看报纸了解相关新闻, 其获取信息的载体是()。
- A. 文字和图像 B. 动画 C. 视频 D. 声音
17. 学生通过观看教学光盘影像进行学习, 那么学生获取信息的载体是()。
- A. 声音 B. 光盘 C. 视频 D. 文字
18. 天气预报可以被多个人接收并且多次使用, 这主要体现信息的()特征。
- A. 时效性 B. 复用性 C. 共享性 D. 传播性

19. 以下有关信息概念的描述, 正确的是()。

A. 信息是一种能量

B. 信息是社会产物

C. 信息区别于资源

D. 信息区别于物质

20. 小明的爸爸说“我的血压是高压150, 低压100, 属于高血压症状”, 这句话描述的是()。

A. 数据

B. 信息

C. 知识

D. 都不是

21. 数据是用于记录客观事物的符号, 其记录形式是多种多样的。以下属于数据的是()。

A. 数学练习册

B. 飞天壁画

C. 移动硬盘

D. 文件袋

22. 以下不属于数据的是()。

A. 课文朗读录音

B. 招生宣传短视频

C. 多功能计算器

D. 卫星气象云图

23. 商品厂家找艺人代言做广告, 以提高产品销售量。这主要体现的信息特征是()。

A. 依附性

B. 价值性

C. 普遍性

D. 可处理性

24. 从身份证号码中可以获得性别信息, 这主要体现的信息特征是()。

A. 时效性

B. 真伪性

C. 共享性

D. 可处理性

25. 以下关于信息概念的描述, 正确的是()。

A. 信息和物质会相互运动

B. 信息一定包含符号

C. 信息与数据会相互运动

D. 信息是事物运动的方式

26. 以下关于信息的描述, 错误的是()。

A. 信息是一种资源

B. 信息是事物及其联系

C. 信息可以表达数据的含义

D. 信息可以减少不确定量

27. 以下关于信息的说法, 正确的是()。

A. 信息可以单独存在, 不需要依附于载体

B. 同一条信息对所有人都具有相同的价值

C. 人类到20世纪才学会存储和处理信息

D. 信息可以被人类感知、处理和利用

28. 数据的记录形式是多种多样的。下列属于数据的是()。

- A. 信息技术教材
- B. 教材配套光盘
- C. 黑板上的文字
- D. LED显示屏

29. 以下有关数据的说法, 错误的是()。

- A. 下课铃声属于非数值型数据
- B. 蔬菜的价格和产量属于数值型数据
- C. 图文声像是数据常见的呈现方式
- D. 计算机能直接识别和处理各种数据

30. 以下关于数据、知识、信息之间关系的表示方法, 正确的是()。

- A. 信息→数据→知识
- B. 数据→知识→信息
- C. 数据→信息→知识
- D. 知识→信息→数据

第二课 数制、数据编码与存储单位

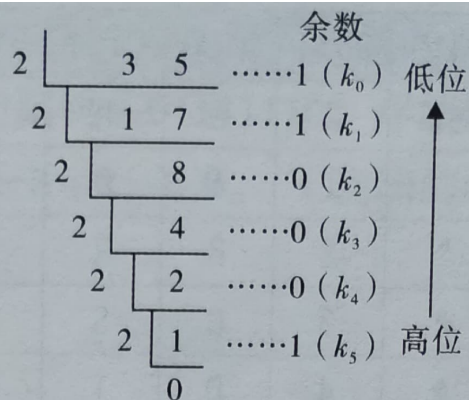
【知识梳理】

一、二进制数

在计算机内部所有的数据都是用二进制数表示的。二进制只有0, 1两个数码能被机器直接识别, 并且运算简单。二进制的运算规则是“逢二进一”“借一当二”。类似地, n 进制的运算规则是“逢 n 进一”“借一当 n ”, 如五进制的进位规则是“逢五进一”“借一当五”。二进制数的表示方式通常是在数的末尾加字母B, 如1101B。

1. 十进制整数化为二进制数

转换方法: 除2取余法(辗转相除), 将欲转化的数反复除以2, 每次取出余数, 商作为下次的被除数, 直至被除数为0停止, 最后将余数按由下而上的顺序写出来即可。例如, 将十进制数35化为二进制数是多少呢? 答案是100011B, 计算过程如下页图所示。



2. 二进制数转换为十进制数

转换方法：采用“按权展开”求和。

例如， $(10111)_2 = 1 \times 2^0 + 1 \times 2^1 + 1 \times 2^2 + 0 \times 2^3 + 1 \times 2^4 = 1 + 2 + 4 + 0 + 16 = (23)_{10}$ 。

小练习：将十进制数27转换为二进制数；将二进制数110101转换为十进制数。

二、信息的编码

1. 信息编码的作用与意义

编码是指按预先规定的方法将文字、数字或其他数据转换为规定的符号组合。在计算机中，编码是指将符号、数字、文字、图像、声音和视频等对象按照预先规定的方法编成二进制代码的过程。编码的作用与意义主要有便于机器识别、存储、传输和处理；方便使用，提高数据的利用率；节省存储空间；提高处理效率等。

2. ASCII 码

ASCII码是“美国国家信息交换标准代码”的简称。它用于对英文字母（大小写不同）、数字0~9、标点符号及控制字符进行编码。ASCII码用一个字节表示一个字符，最高位固定为0，其实只用到7位，总共能表示 2^7 （等于128）个字符。在计算机内部ASCII码用二进制数表示，使用时也可以转化为用等值的十进制数表示。部分字符的ASCII如下表所示。查阅某个字符ASCII码的方法是先找到对应列顶端的高四位，再找到对应行左边的低四位，二者组合在一起即可，如“A”的ASCII码是01000001。

二进制	0010	0011	0100	0101	0110	0111
0000	空格	0	@	P	,	p
0001	!	1	A	Q	a	q
0010	"	2	B	R	b	r
0011	#	3	C	S	c	s
0100	\$	4	D	T	d	t
0101	%	5	E	U	e	u
0110	&	6	F	V	f	v
0111	'	7	G	W	g	w
1000	(	8	H	X	h	x
1001	)	9	I	Y	i	y
1010	*	:	J	Z	j	z

常见字符ASCII码（十进制表示）的范围如下表所示。

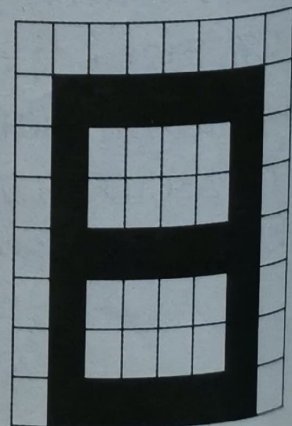
ASCII 码（十进制）	对应的字符
0 ~ 31	控制字符
32	空格
48 ~ 57	数字 0 ~ 9
65 ~ 90	大写字母 A ~ Z
97 ~ 122	小写字母 a ~ z
123 ~ 128	英文标点符号、运算符

ASCII码的大小关系：空格<数字<大写字母<小写字母。

学会根据已知字符的ASCII码，推导其他字符的ASCII码。例如，已知“d”的十进制ASCII码为100，那么“g”的十进制ASCII码为103。推算方法：“g”与“d”相差3， $100+3=103$ 。

3. 汉字的编码

处理汉字的不同环节使用不同的编码，汉字从输入到输出各环节的编码是输入码→机内码→国标码（GB2312）→输出码（字形码）。存储一个汉字的机内码或国标码用2个字节，国标码每个字节的最高位都是0，如存储1000个汉字占用的存储空间大约是2 KB。汉字的输出码用字形点阵表示，如右图所示是“日”字的



8×8点阵码,假设空白方块用“0”表示,黑色方块用“1”表示,则从上往下第二行,从左往右的编码是01111110。存储一个如上页图所示的字形点阵占用的存储空间是 $8 \times 8 \div 8 = 8$ B。(位图的编码类似于汉字的字形点阵编码。)

三、信息存储单位

存储信息的最小单位是位(bit,简记为b,即比特),基本单位是字节(Byte,简记为B), $1\text{ B} = 8\text{ bit}$ 。此外还有千字节(KB),兆字节(MB),吉字节(GB),太字节(TB)等。 $1\text{ KB} = 1024\text{ B}$, $1\text{ MB} = 1024^2\text{ B} = 1024\text{ KB}$, $1\text{ GB} = 1024^3\text{ B} = 1024\text{ MB}$, $1\text{ TB} = 1024^4\text{ B} = 1024\text{ GB}$ 。

【单元练习】

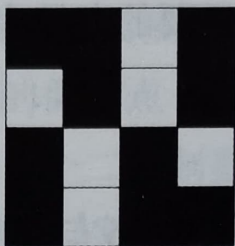
- 将二进制数 $(1011)_2$ 转换为十进制数等于()。
A. 10 B. 11 C. 12 D. 13
- 与十进制数14等值的二进制数是()。
A. 1110 B. 1101 C. 1100 D. 1111
- 以下二进制数中最大的一个是()。
A. 11100 B. 1111 C. 10111 D. 11011
- 二进制数的表示方法通常是在数值的后面加上字母()。
A. D B. B C. O D. H
- 将二进制数10101转换为十进制数,下列正确的是()。
A. 42 B. 22 C. 21 D. 43
- 通常将十进制整数转换为R(R为2, 8, 16)进制数采用的方法是()。
A. 按权展开求和法 B. 拆数法
C. 移位法 D. 除R取余法
- 以下可能是二进制数的一项是()。
A. 1456 B. 1201 C. 98AF D. 1011
- 下列将十进制数29转换为二进制数,正确的是()。
A. 10110 B. 11001 C. 11101 D. 10111

9. 数值数据通常采用数制来表达。以下关于四进制数进位规则的描述, 正确的是 ()。
- A. 逢二进一 B. 逢十进一 C. 逢一进四 D. 逢四进一
10. 4位二进制数的组合状态数是 ()。
- A. 4 B. 8 C. 16 D. 12
11. 在计算机中存储信息的最小单位是 ()。
- A. 字节 B. 二进制位 C. 千字节 D. 兆字节
12. 在计算机中存储信息的基本单位是 ()。
- A. 比特 B. KB C. MB D. 字节
13. 1 MB等于 ()。
- A. 1000 KB B. 1024 B C. 1000×1000 B D. 1024×1024 B
14. 以下对信息存储容量单位的换算, 错误的是 ()。
- A. 1 TB=1024 GB B. 1 MB=1024 KB C. 1 KB=1024 MB D. 1 GB=1024 MB
15. 一个字节相当于 ()。
- A. 1 bit B. 8 bit C. 2 bit D. 16 bit
16. 以下对存储容量单位大小的比较, 正确的是 ()。
- A. 1 GB>1 MB>1 TB B. 1 TB>1 MB>1 KB C. 1 GB>1 TB>1 MB D. 1 MB>1 KB>1 TB
17. 在微型机中, 英文字符的编码方案是 ()。
- A. BCD码 B. 区位码 C. ASCII码 D. 拼音码
18. 7位ASCII码所能表示的字符个数是 ()。
- A. 64 B. 128 C. 256 D. 1024
19. 已知字符“A”的ASCII码为01000001, 那么字符“C”的ASCII码为 ()。
- A. 01000010 B. 01000101 C. 01000100 D. 01000011
20. 以下字符中, ASCII值最大的是 ()。
- A. z B. D C. 9 D. 空格

21. 若字符“B”的ASCII码为66, 那么字符“F”的ASCII码为()。

- A. 68 B. 69 C. 70 D. 71

22. 下图为4×4像素组成的位图, 按从左到右的编码规则, 第一行的编码是1101, 则第二行的编码是()。



- A. 0011 B. 0101 C. 1100 D. 1010

23. 已知“g”的ASCII码值的十进制数是103, 则“d”的ASCII码值的十进制数是()。

- A. 99 B. 101 C. 100 D. 98

24. 以下有关ASCII码的说法, 错误的是()。

- A. ASCII码是美国信息交换标准码的简称
B. ASCII码主要用于对字母、数字和标点符号的编码
C. ASCII码是汉字的统一编码方案
D. 一个字符的ASCII码占一个字节, 最高位固定为0

25. 下表为部分字符的ASCII码, 其中字符“#”的ASCII码是()。

二进制	0000	0001	0010	0011	0100	0101	0110	0111
0000	NUL	DLE	空格	0	@	P	,	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	"	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	EOT	DC4	\$	4	D	T	d	t
0101	ENQ	NAK	%	5	E	U	e	u

- A. 00100011 B. 00110010 C. 01000011 D. 00100101

26. 1 KB存储容量能存储的英文字符个数是()。

- A. 128个 B. 1024个 C. 512个 D. 2048个

27. 1 KB的存储容量最多能存储汉字的个数是()。

- A. 1024 B. 512 C. 2048 D. 1000

28. 存储1024个汉字的国标码占用的存储空间是()。

- A. 512 B B. 1 KB C. 2 KB D. 128 B

29. 汉字在机器内部进行存储、加工和传输使用的代码是()。

- A. 机内码 B. 拼音码 C. 自然码 D. 字型码

30. 存储10 000个英文字母需要的存储容量大约是()。

- A. 5 KB B. 2 KB C. 10 KB D. 1 MB

第三课 声音与图像编码

【知识梳理】

一、模拟信号与数字信号

1. 模拟信号

模拟信号是指用连续变化的物理量所表达的信息，相对不稳定，在传播的过程中容易受到干扰而导致失真。用传感器直接获得的信号，如声音、图像、视频、温度、压强等都是模拟信号。模拟信号的值叫模拟数据。

2. 数字信号

数字信号是离散时间信号的数字化表示。数字信号的值称为数字数据，数字数据用二进制数0和1表示。模拟信号与数字信号可以相互转换。

二、模拟数据数字化的过程

将声音、图像等模拟数据数字化过程的基本步骤：采样→量化→编码。

1. 声音的数字化

声音的采样是指每隔一段时间在模拟声音信号的波形上取一个幅度值，1秒内的采样次数称为采样频率，采样频率越高，声音的保真度越高。声音的量化是指用二进制数表示采样所得到的幅度值的过程，量化位数越大，声音的保真度越高。将量化后的数据转化为二进制数的过程称为编码。

对于音频信号来说，采样频率越高，量化位数越多，声音的保真度越高，声音的质量越好，但所需存储空间越大。

声音数据容量(B) = (采样频率 × 量化位数 × 声道数 × 持续时间) ÷ 8。

例如，录制一段采样频率为22.05 kHz、量化位数为16位、时长为10秒的双声道立体声音频，其文件的存储容量 $=22.05 \times 1000 \times 16 \times 2 \times 10 \div 8 = 882\,000$ (B)， $882\,000 \div 1024 \approx 861.3$ (KB)。

2. 图像的数字化

图像的采样是在空间上按一定间隔从左到右、自上而下提取画面信息，将图像分割成若干个离散的像素点；图像的量化，就是将采样得到的每个像素点的颜色或亮度用若干个二进制数表示出来；图像的编码就是按一定的格式把位图上的各个像素点的量化数据记录下来的过程。

图像容量 (B) = 分辨率 $\times$ 颜色位深度 $\div 8$ 。例如，存储一幅大小为 20×30 像素，颜色深度为24位的未压缩图像，其容量为 $20 \times 30 \times 24 \div 8 = 1800$ (B)。

颜色种数 $=2^n$ (n 表示颜色的位数)。

一个二进制位有两种状态，如用1表示黑色，0表示白色，一个二进制位可以表示任意的两色，那么 n 个二进制位就有 2^n 种组合，可以表示最多 2^n 种颜色，如6个二进制位可表示 $2^6=64$ 种颜色。例如，存储一幅大小为 20×30 像素，颜色为16色的图像， $2^4=16$ ，那么每个像素颜色位数是4位，图像容量为 $20 \times 30 \times 4 \div 8 = 300$ (B)。

3. 图像的分类

根据生成图像的原理不同，图像分为位图（点阵图）和矢量图（向量图）。位图由像素组成，其缺点是放大容易失真，文件容量大，优点是色彩丰富，能逼真地反映自然景物；矢量图是由线条加色块组成，其优点是放大不容易失真，文件容量小，缺点是色彩单调不够丰富。

三、文件扩展名

文件扩展名用于表示文件的类型或格式，常用的文件扩展名如下：

(1) 文本文件扩展名：txt（记事本文件），doc/docx（Word文档），ppt/pptx（PowerPoint演示文稿），wps（WPS文档）等。

(2) 声音文件扩展名：wav，mp3，wma，mid，au，ra，其中wav格式是存储原始录音的波形声音，音质好，但未经压缩，文件容量大。

(3) 图像文件扩展名：jpg/jpeg，gif，bmp，png，tif/tiff，psd等。

(4) 视频文件扩展名: avi, mpg, wmv, mp4, asf, mov等。

(5) 其他格式的文件扩展名: rar/zip (压缩文件), xls/xlsx (Excel 电子表格) 等。

【单元练习】

1. 随着时间连续变化的数据称为 ()。

A. 数值数据 B. 动态数据 C. 数字数据 D. 模拟数据

2. 随时间变化是离散的数据称为 ()。

A. 间接数据 B. 数字数据 C. 模拟数据 D. 可变数据

3. 以下对数据的叙述, 错误的是 ()。

A. 数字数据不能转化为模拟数据
B. 在计算机中, 数据以二进制的形式表示
C. 文字、图像、声音、视频等数据属于模拟数据
D. 模拟数据必须转化为数字数据, 计算机才能加工处理

4. 声音的数字化过程中, 正确的步骤顺序是 ()。

①量化 ②采样 ③编码

A. ①②③ B. ②①③ C. ②③① D. ①③②

5. 录音设备在1秒内对声音信号的采样次数称为 ()。

A. 采样精度 B. 采样频率
C. 采样样本 D. 采样数量

6. 采用以下标准采集的音频, 质量最高的是 ()。

A. 采样频率为22.05 kHz, 量化位数为8位
B. 采样频率为22.05 kHz, 量化位数为16位
C. 采样频率为44.1 kHz, 量化位数为8位
D. 采样频率为88.2 kHz, 量化位数为16位

7. 声音采集和量化过程中, 采样频率越高, 量化位数越多, 则 ()。

①声音的保真度越高 ②音质效果越差 ③数据量越大 ④存储容量越小

A. ①③

B. ②④

C. ①②③

D. ①②④

8. 将模拟数据转换为数字数据的过程包含编码和()。
- A. 量化、采样 B. 压缩、量化
C. 转换、采样 D. 解密、量化
9. 用0, 1表示的数据称为()。
- A. 模拟数据 B. 物理数据 C. 模拟信号 D. 数字数据
10. 计算机要处理声音必须配备的设备是()。
- A. 网卡 B. 声卡 C. 显示卡 D. 视频采集卡
11. 存储一幅尺寸和分辨率相同的图像, 以下容量最小的是()。
- A. 16色图像 B. 黑白图像
C. 256色图像 D. 24位真彩色图像
12. 与图像容量大小密切相关的因素有()。
- ①图像尺寸 ②分辨率 ③图像制作软件 ④颜色位数
- A. ①②④ B. ①③④ C. ②③④ D. ①②③
13. 存储一幅未经压缩、分辨率为 10×20 、颜色为16色的位图图像, 占用的存储空间是()。
- A. 3 KB B. 400 B C. 200 B D. 100 B
14. 存储一幅未经压缩、分辨率为 20×10 、颜色位数为24位的位图图像, 其存储容量为()。
- A. 800 B B. 600 B C. 4800 B D. 480 B
15. 5位二进制数最多能表示的颜色数为()。
- A. 16 B. 8 C. 32 D. 4
16. 以下不属于音频文件格式的是()。
- A. mp3 B. mov C. wav D. wma
17. 以下属于文本文件的是()。
- A. 校园风采.avi B. 校园风采.gif
C. 校园风采.txt D. 校园风采.xls
18. 下列属于视频文件的是()。
- A. xysp.zip B. xysp.mid C. xysp.wmv D. xysp.pdf
19. 以下不属于图像文件格式的是()。
- A. doc B. png C. jpg D. bmp

20. 某位图图像包含256种颜色, 此图像的色彩位数至少为()。
- A. 4位 B. 16位 C. 32位 D. 8位
21. 利用格式转换软件将原始的课文朗读录音“春.wav”转换为“春.mp3”, 将会()。
- A. 提高音质 B. 缩小文件的存储容量
C. 增大文件的存储容量 D. 降低噪声
22. 将一幅bmp格式的图像转换为jpg格式, 则()。
- A. 文件的存储容量变小 B. 文件的存储容量变大
C. 色彩变丰富 D. 尺寸变大
23. 小华从网上下载一个视频, 其文件格式不可能是()。
- A. mp4 B. avi C. ppt D. mpg
24. 采用以下指标录制一段相同时长的立体声音频, 其文件容量最大的是()。
- A. 采样频率为22.05 kHz、量化位数为8位
B. 采样频率为44.1 kHz、量化位数为8位
C. 采样频率为48.0 kHz、量化位数为16位
D. 采样频率为88.2 kHz、量化位数为16位
25. 采集一段采样频率为22.05 kHz、量化位数为16位、时长为5秒的双声道立体声音频, 其文件的存储容量大约为()。
- A. 88.2 KB B. 3528 KB C. 220.5 KB D. 430.7 KB

专题二 数据处理与应用

【学习目标】

1. 了解数据的价值与数据处理的意义。
2. 了解数据处理的基本过程。
3. 掌握数据收集与整理的方法。
4. 掌握数据采集工具的使用方法。
5. 体验数字化学习，感受利用数字化工具和资源的优势。
6. 了解数据库管理系统的基本知识。
7. 了解数据分析的常用方法，能使用电子表格软件进行简单的数据分析。
8. 掌握常见的图表类型及其特点，能依据需求选择适当的图表进行可视化呈现。
9. 能对整个数据分析过程进行总结，并以报告的形式呈现出来，为决策者提供参考。
10. 理解对数据进行保护的意义。

第一课 数据采集与处理

【知识梳理】

一、数据处理与应用

数据处理的一般过程是定位数据需求、收集数据、加工处理、分析数据和撰写分析报告、使用数据。数据处理过程的主要环节包括采集、存储、传输、加工、检索、分析和输出等。

以网上阅卷系统为例，扫描试卷（采集数据）、保存扫描后的试卷和评分结果（存储数据）、分发试卷给阅卷老师和上传阅卷结果（传输数据）、客观题自动阅卷和统分（加工数据）、老师按姓名或考号查看学生答题情况（检索数据）、系统自动分析学生的进步情况（分析数据）、打印考试分析结果（输出数据）。

1. 数据的来源与采集方法

按数据的获取渠道分为直接数据（一手数据）和间接数据（二手数据）。直接数据的获取方法有调查法、访谈法、观察法、实验法和利用传感器采集；间接数据的获取方法有他人告知，通过书报刊、互联网、广播与电视等媒体进行采集。

2. 数据来源可信度高的渠道

数据来源可信度高的渠道有政府部门、权威媒体、官方网站等。

3. 数据的采集工具与方法

（1）文字采集方法。键盘输入、语音识别成文本、手写输入、光学字符识别（OCR）、用手机拍图识字APP将纸质文本输入。

（2）图像的采集方法。利用智能手机、上网本、数码相机拍照，纸质图像扫描输入，电脑截图、截屏；网上下载，软件制作等。

（3）声音的采集方法。用话筒、麦克风、录音笔、录音机录制，网上下载，从已有的音频中剪辑等。

（4）视频的采集方法。用智能手机、笔记本电脑、摄像头、数码摄像机拍摄；录屏，网上下载，从已有的视频中剪辑，从电视节目中录制等。

4. 数据分析的基本方法

数据分析的方法很多，如对比分析、平均分析、交叉分析等。对比分析是将两个或两个以上的数据进行比较，分析它们的差异，从而揭示这些数据所代表的事物的发展变化情况和规律，如为了减少交通事故的发生，交通管理部门对事故多发地段历年来的数据进行对比，改善交通措施。平均分析是运用计算平均值的方法反映总体在一定时间、地点条件下某一数量特征的一般水平，如对学生的考试成绩取平均分进行分析。交叉分析是从横向和纵向两个方向来计算两个或多个有联系的变量在交叉点的统计值。

5. 数据的可视化呈现

数据可视化是指以图形、图表、动画等方式更加直观生动地呈现数据及数据分析的结果，以便揭示数据之间的关系、变化趋势和规律等表达方式。常见的可视化工具有思维导图、统计图表、导航地图、词云图、标签

云图、流程图、E-R图等。常见的统计图表有柱形图、条形图、折线图、饼图、散点图、面积图、雷达图和股价图等。

词云也叫文字云，是对文本中出现频率较高的“关键词”予以视觉化的展现。词云图过滤掉大量低频的文本信息，使浏览者一眼就能看出表达的重点。例如，通过右边的

“福建旅游”词云图，我们可以很直观地看出福建有哪些主要旅游景点。



二、数字化工具的使用

凡是利用电脑、网络、智能手机及各种数码工具进行的信息活动都属于数字化工具的使用。使用数字化工具与资源的优势有共享性、便捷性、互动性、高效性、多样性等。数字化工具给我们的学习和生活带来极大的便利，例如：我们可以借助多媒体课件进行自主学习；利用微信交流学习、传输文件；利用电子邮件提交作业和材料。

三、数据库管理系统（了解性）

数据库（简称DB）是数据的仓库，利用数据库来管理大批量的数据便于数据共享、检索、更新和维护等。常见的数据库软件有FoxPro、Access、SQL、Oracle、Sybase等，数据库软件的核心是数据库管理系统（DBMS），利用数据库软件设计的面向实际应用的信息系统称为数据库应用系统，如图书管理系统、选课管理系统、人事管理系统和高考成绩查询系统等。

四、数据保护的措施

数据安全保护是指数据不被破坏、更改、泄露或丢失。保护数据安全的措施有安装杀毒软件、安装防火墙、数据备份、访问控制、数据加密等。在大数据时代，数据共享和传播速度的加快，更容易造成商业信息和机密信息的泄露，数据发布者需要加强自身的保护意识，并借助一定的技术手段防范数据泄露，同时也需要对数据使用者进行道德和法律上的约束。

设置密码的原则：密码长度越长强度越高；尽量使用字母（大小写不同）、数字和特殊符号相结合的密码；不要使用电话号码、身份证号或出

【单元练习】

1. 高考网上阅卷之前,需要先扫描试卷,这个过程属于()。
A. 数据采集 B. 数据处理 C. 数据输出 D. 数据加密
2. 数据处理过程包括数据采集、传输、加工、存储、检索和()。
A. 删除 B. 输出 C. 评价 D. 转换
3. 以下数据来源中,属于“一手数据”的是()。
A. 听专家讲座 B. 微信朋友圈 C. 网络搜索 D. 科学实验
4. 下列对信息来源的叙述,错误的是()。
A. 在收集信息时,信息来源越多越好
B. 在收集信息时,应该合理选择信息来源
C. 信息源越广阔,收集到的信息量就越大
D. 要力求信息来源具有一定的代表性
5. 以下属于间接获取数据的渠道是()。
A. 事件现场录音 B. 现场调查
C. 浏览报纸 D. 测量体温
6. 以下属于收集“直接数据”的方法是()。
A. 观察法 B. 查阅文献资料
C. 教学光盘 D. 询问老师
7. 同学们要获取“学校发展史”的数据,以下最有效的获取途径是()。
A. 收听广播 B. 百度搜索
C. 观看电视节目 D. 到校史室查阅档案
8. 要获取某大学计划招生人数的数据,以下信息来源中可信度最高的是()。
A. 招生代理机构 B. 网友告之
C. 该大学的官方招生网站 D. 网络论坛
9. 以下不属于网络数据获取方法的是()。

- A. 使用搜索引擎 B. 网络论坛
- C. 拨打客服电话 D. 网络数据库
10. 利用百度搜索信息时, 必须在搜索框中输入()。
- A. IP地址 B. 关键词 C. 域名 D. 邮件地址
11. 乘坐飞机时通过人脸识别获得旅客信息, 这个过程属于()。
- A. 数据加工 B. 数据采集
- C. 数据整理 D. 数据通信
12. 以下不属于数据库管理系统的是()。
- A. Excel 2003 B. Access
- C. SQL D. Foxpro
13. 以下属于通过数字化工具获取信息的是()。
- A. 拨打电话获得天气情况
- B. 当面询问老师获取成绩
- C. 使用纸质问卷调查了解学生健康信息
- D. 使用手机拍摄获取班级活动照片
14. 在“12306”网站查询动车票, 其实查询的是()。
- A. 网页文件 B. 电子表格 C. 数据库 D. 文本文件
15. 某餐厅利用在线点菜系统管理菜单, 该系统属于()。
- A. 数据库 B. 人工管理系统
- C. 数据库应用系统 D. 数据库管理系统
16. 为了预防流行性疾病, 做好防范措施, 卫生健康部门对历年的数据进行对比。这属于数据处理过程中的()。
- A. 数据查找 B. 数据输出
- C. 数据拆分 D. 数据分析
17. 以下软件可以用于处理和分析数据的是()。
- A. Excel B. Photoshop C. Outlook D. Flash
18. 期末考试结束后, 年段长对各科成绩取平均分进行考情分析, 这种分析法属于()。
- A. 对比分析 B. 交叉分析 C. 平均分析 D. 综合分析
19. 要录制一段校长在升旗仪式上的讲话的音频, 不可以使用()。

- A. 录音笔 B. 扬声器 C. 录音机 D. 智能手机

20. 以下设备可以用于采集视频的有()。

- ①智能手机 ②投影仪 ③数码相机 ④笔记本电脑 ⑤传真机
A. ①②③ B. ①②④ C. ①③④ D. ③④⑤

21. 以下设备均可用于获取数字化图像的是()。

- A. 录音笔、投影仪 B. 数码相机、智能手机
C. 数码相机、打印机 D. 数码摄像机、手写板

22. 将漫画书上的图片保存到计算机中, 以下可用的数字化采集工具是()。

- A. 打印机 B. 绘图仪 C. 扫描仪 D. 投影仪

23. 为了做好校园歌手比赛的图文报道, 以下合适的信息采集设备是()。

- A. 扫描仪 B. 打印机 C. 数码相机 D. 投影仪

24. 采集一段文本信息不可以使用()。

- A. 键盘 B. 手写板 C. 扫描仪 D. 打印机

25. 铁观音茶文化研究性学习小组为了更好地开展研究工作, 除了上网查资料外, 他们还拟订了获取第一手资料的其他方法, 其中比较便捷的是()。

①考察学校附近的茶厂 ②访问学校附近的茶商 ③收听广播 ④亲自品茶

- A. ①②④ B. ②③④ C. ①②③ D. ①③④

26. 以下学习方式, 没有用到数字化工具的是()。

- A. 当面请教老师解题方法 B. 浏览专家博客学习写作方法
C. 与同学在微信上讨论数学难题 D. 借助百度词典学习英语单词

27. 以下不属于数据可视化表达方式的是()。

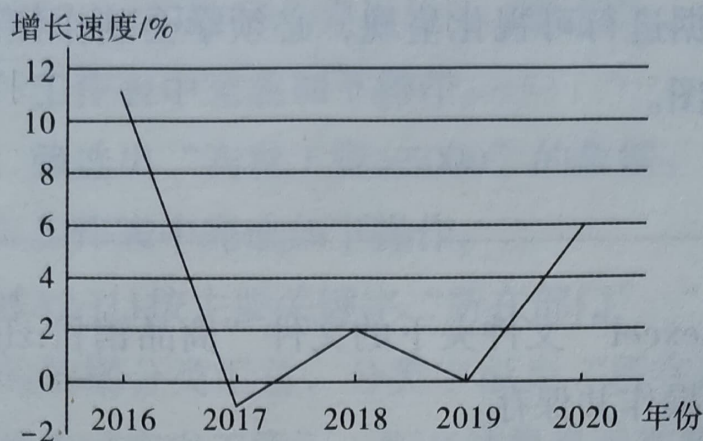
- A. 统计图表 B. 模拟实验动画
C. 语音助手 D. 标签云图

28. 以下可实现实时数据可视化的是()。

- A. 流程图 B. 导航地图 C. 思维导图 D. 柱形图

29. 下图为统计局网站呈现的某种数据图表。这种呈现数据的形式称

为()。



A. 文字化形式

B. 可视化形式

C. 多样化形式

D. 虚拟化形式

30. 以下选项中,不属于数据安全防护技术的是()。

A. 数据备份

B. 数据加密

C. 安装防火墙

D. 数据共享

第二课 Excel基本知识

【知识梳理】

一、Excel 的基本知识

1. 工作表的插入、改名、复制、移动和删除。
2. 文本格式的设置: 字体、字号、字形及颜色设置。
3. 单元格“合并及居中”, 数据类型及小数位数设置。
4. 对齐方式与边框的设置(“所有框线”)。
5. 单元格行高与列宽的设置。
6. 单元格区域中数据序列的填充。
7. 数据计算(公式法和函数法), 公式以“=”开头; 函数有求和sum(), 求平均average(), 求最大值max(), 求最小值min()。
8. 数据排序, 根据一个或两个关键字排序, 排序时一般是选择全部数据(标题也要选择)。

(备注: 本书中使用的Excel为2003版。)

二、图表制作

制作图表对数据进行可视化呈现，必须学会制作柱形图、条形图、折线图、饼图和散点图。

【单元练习】

1. 打开“test\excel”文件夹下的文件“商品销售.xls”，在“Sheet1”工作表中完成如下操作并保存。

(1) 将单元格区域A1:E1“合并及居中”，设置字体为“黑体”，字号为“16”，颜色为“红色”。

(2) 完成单元格区域A4:A8“商品编号”的自动填充。

(3) 在单元格区域E3:E8中，使用公式法分别计算各商品的总价，总价=数量×单价。

(4) 将单元格区域E3:E8的数字分类设置为“货币”，货币符号为“¥”，小数位数为1位。

(5) 在单元格C9和D9中，使用函数法分别统计单价和数量的最大值。

(6) 在单元格C10和D10中，使用函数法分别统计单价和数量的平均值。

(7) 将单元格区域A2:E10的边框设置为“所有框线”。

(8) 将第2~10行的行高设为20，第A~E列的列宽设置为8。

(9) 将单元格区域A2:E2的背景填充为标准色“橙色”。

(10) 将单元格区域A2:E10的水平对齐方式设置为“居中”。

(11) 删除“Sheet2”工作表。

(12) 将当前工作表重命名为“销售汇总”，并复制当前工作表，放在当前工作表的后面。

2. 打开“test\excel”文件夹下的文件“工资统计.xls”，按照以下要求操作并保存。

(1) 在“表1”工作表中完成如下操作。

①将标题的水平对齐方式设置为“居中”。

②在单元格区域I3:I11中，使用公式法分别计算每个员工的实发工资，实发工资=基本工资+奖金+应扣工资。

③将单元格区域A2:I11的外边框设置为双实线，颜色为“红色”，内边

框设置为细单实线。

④将单元格区域F3:I11的数字分类设置为“数值”，小数位数为1位。

(2) 在“表2”工作表中完成如下操作。

利用自动筛选，筛选出“实发工资>5000”的数据。

(3) 在“表3”工作表中完成如下操作。

①将单元格区域A2:I11按主要关键字“所在部门”“升序”排序。

②将工作表中的数据分类汇总，分类字段为“所在部门”，汇总方式为“求和”，汇总项为“实发工资”，汇总结果显示在数据表下方。

3. 打开“test\excel”文件夹下的文件“家庭支出.xls”，按照以下要求操作并保存。

(1) 在“Sheet1”工作表中完成以下操作。

①利用填充柄完成单元格区域A3:A14中月份的填充。

②在H3:H14单元格区域中，使用函数法分别计算各月份的总支出。

③交换“房租”和“水电”两列数据。

④选择单元格区域A2:H14，按主要关键字“总支出”“降序”排序，次要关键字“医疗”“升序”排序。

⑤选择单元格区域A2:B14，插入一个“簇状柱形图”，系列产生在“行”。设置图表标题为“水电费统计图”；在“图例”标签中，取消“图例”的选择；在“数据标志”标签中，选择“值”选项。

(2) 在“Sheet2”工作表中完成如下操作。

选择单元格区域A2:A14和H2:H14，插入一个“簇状条形图”，系列产生在“行”。设置图表标题为“总支出统计图”，双击图表标题，在“字体”标签中，选择“双下划线”、颜色“蓝色”；在“坐标轴”标签中，取消“分类(x)轴”的选择。

(3) 在“Sheet3”工作表中完成如下操作。

选择单元格区域A2:G3，插入一个“分离型饼图”，系列产生在“行”。设置图表标题为“1月各项支出占比”，在“数据标志”标签中，选择“百分比”选项。

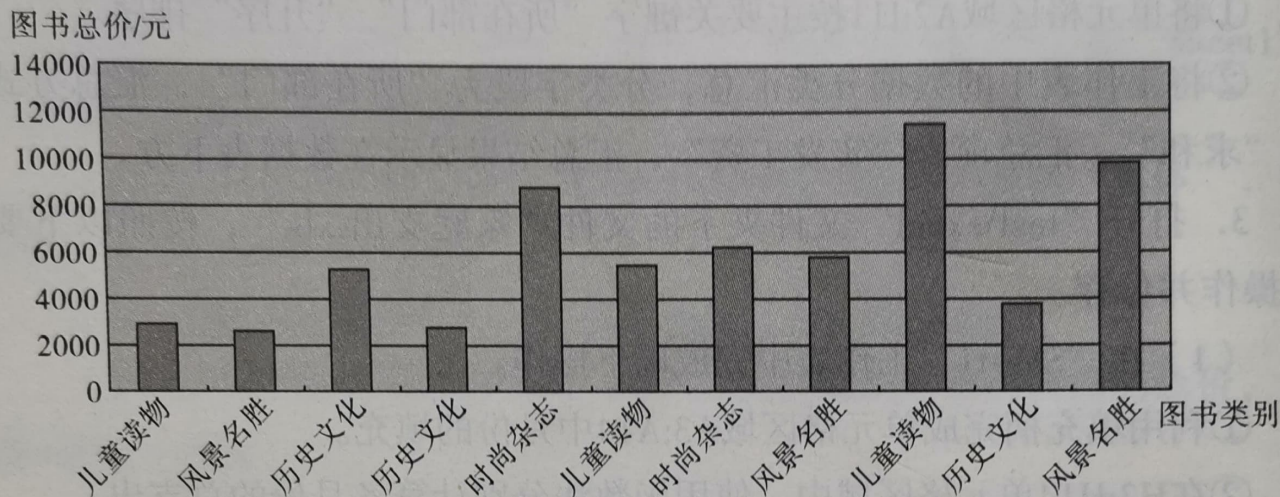
(4) 在“Sheet4”工作表中完成如下操作。

选择单元格区域A2:A14和H2:H14，插入一个“折线图”，系列产生在

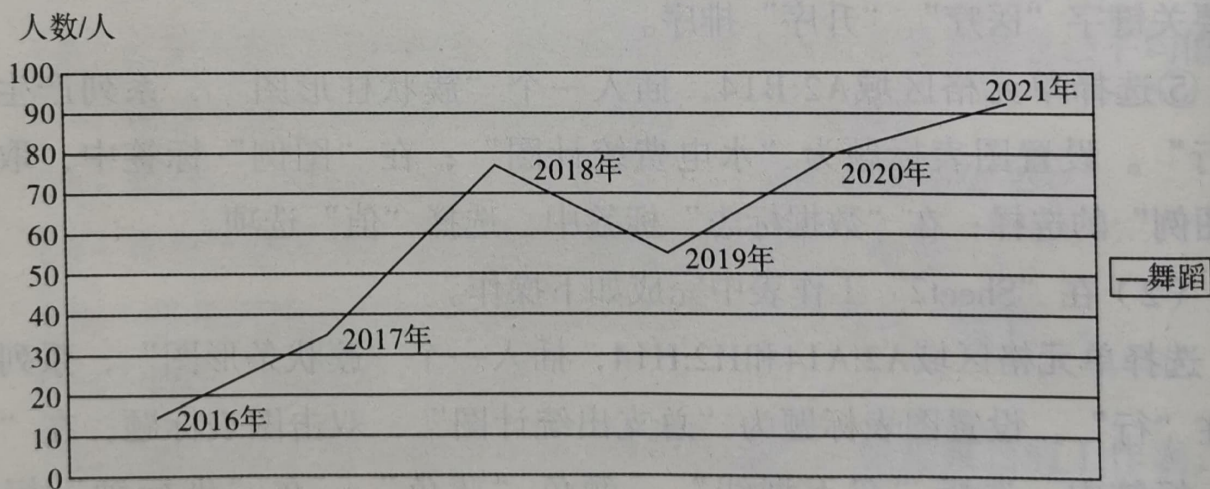
“列”。设置图表标题为“1月至12月总支出变化图”，在“图例”标签中，选择位置为“靠左”。

4. 根据图表样式制作图表，打开“test\excel”文件夹下的文件“图表制作素材.xls”，按以下要求制作图表。

(1) 在“Sheet1”工作表中，制作一个如下图所示的图表。



(2) 在“Sheet2”工作表中，制作一个如下图所示的图表。



(3) 在“Sheet3”工作表中，制作一个如下图所示的图表。

第4季度各种消费品销售比例

