

- اس باب کے اختتام تک، آپ اس قابل ہو جائیں گے کہ:
- ڈیٹا سائنس کے تصور اور روزمرہ زندگی میں اس کی اہمیت کی وضاحت کر سکیں۔
- ڈیٹا سائنس لائف سائیکل (DSL) کے ہر مرحلہ کو بیان کر سکیں۔
- حقیقی دنیا کے ایسے مسائل کی شناخت اور تعریف کر سکیں جنہیں ڈیٹا کا استعمال کرتے ہوئے حل کیا جاسکتا ہے۔
- سروے، آن لائن فارمز، اور مشاہداتی تکنیک کا استعمال کرتے ہوئے ڈیٹا اکٹھا کر سکیں۔
- درستگی (accuracy) کو بہتر بنانے کے لیے ڈیٹا کلیٹنگ اور تصدیق (validation) کا عمل انجام دے سکیں۔
- ڈیٹا بیس سے ڈیٹا محفوظ کرنے اور بازیافت کرنے کے لیے سادہ SQL سوالات (queries) (جیسے INSERT, SELECT) کا استعمال کر سکیں۔
- بنیادی انکوائری (جیسے ترتیب (sorting)، درجہ بندی (classification)، کلسٹرنگ کا استعمال کرتے ہوئے ڈیٹا کا تجزیہ کر سکیں۔
- ڈیٹا سے حاصل ہونے والے بصیرت (insights) کو بصری (visual) طور پر ظاہر کرنے کے لیے بار چارٹس اور ڈیش بورڈز بنا سکیں۔
- چارٹس اور سریز کا استعمال کرتے ہوئے نتائج کو بیان کر سکیں۔

انشائی طرز سوالات

محکمہ تعلیم کی نئی امتحانی تکنیکس (Knowledge, Understanding, Application, Analytical & Conceptual) کی روشنی میں مرتب کیے گئے انشائی طرز سوالات

تعارف

سوال: ڈیٹا سائنس کیا ہے؟ یہ روزمرہ کی زندگی میں ہمارے لیے کس قدر ضروری ہے؟
جواب: ڈیٹا سائنس:

ڈیٹا سائنس مسائل کو سمجھنے، پیشران (نمونوں) کو دریافت کرنے، اور بہتر فیصلے کرنے کے لیے ڈیٹا استعمال کرنے کا عمل ہے۔ موجودہ ڈیجیٹل دور میں ڈیٹا ہر جگہ موجود ہے۔ یعنی سروے سے لے کر موبائل ایپس، سوشل میڈیا، آن لائن شاپنگ، اور صحت کی خدمات تک۔ اہمیت: ڈیٹا سائنس ہمیں اس خام ڈیٹا کو مفید معلومات میں تبدیل کرنے میں مدد دیتی ہے۔ چاہے موسم کی پیش گوئی کرنا ہو، طلبہ کی دلچسپیوں کو سمجھنا ہو، یا صحت کی سہولیات کو بہتر بنانا ہو۔ ڈیٹا سائنس ہمارے جینے، سیکھنے اور کام کرنے کے طریقے کو بدل رہی ہے۔ پوری دنیا میں تنظیمیں فیصلہ سازی کو بہتر بنانے، اخراجات بچانے، اور بہتر خدمات تخلیق کرنے کے لیے ڈیٹا کا استعمال کرتی ہیں۔ یہ باب ڈیٹا سائنس لائف سائیکل (DSL) متعارف کراتا ہے، جس میں مسئلہ کو سمجھنا، ڈیٹا اکٹھا اور صاف کرنا، ڈیٹا کا تجزیہ اور تشریح کرنا، اور آخر میں نتائج کو چارٹس اور بصری مواد کے ذریعے پیش کرنا شامل ہے۔ ہر مرحلہ تنقیدی سوچ اور مسئلہ حل کرنے کی مہارتیں پیدا کرتا ہے۔ طلبہ یہ بھی سیکھیں گے کہ سروے، اسپرڈ شیٹس اور چارٹس جیسے سادہ اوزاروں کا استعمال کرتے ہوئے ڈیٹا کیسے اکٹھا، صاف، تجزیہ، تشریح اور تصور کیا جاتا ہے۔ اس باب میں حقیقی دنیا کی مثالیں اور عملی سرگرمیاں شامل ہیں جو طلبہ کو ڈیٹا سائنس کو کلاس روم اور روزمرہ کی زندگی میں لاگو کرنے میں مدد دیتی ہیں۔

سوال 1: ڈیٹا سائنس کی تعریف کریں اس کی اہمیت بھی بیان کریں؟
جواب: ڈیٹا سائنس مسائل کو سمجھنے، پیٹرن (نمونوں) کو دریافت کرنے، اور بہتر فیصلے کرنے کے لیے ڈیٹا استعمال کرنے کا عمل ہے۔ چاہے ہم موبائل فون استعمال کر رہے ہوں، گوگل پر سرچ کر رہے ہوں، آن لائن شاپنگ کر رہے ہوں، یا پھر کسی انٹرنیٹ پر کچھ تلاش کر رہے ہوں۔ ڈیٹا اکٹھا کیا جا رہا ہوتا ہے۔

ڈیٹا سائنس کا مقصد: ڈیٹا سائنس وہ شعبہ ہے جہاں ہم ڈیٹا کو پیٹرن تلاش کرنے، سوالوں کے جواب دینے، اور لوگوں یا کاروباروں کو بہتر فیصلے کرنے میں مدد دینے کے لیے استعمال کرتے ہیں۔ اس میں ڈیٹا کو اکٹھا کرنا، منظم کرنا، تجزیہ کرنا اور ویژولائزیشن (visualize) کرنا شامل ہے۔
فیصلہ سازی میں ڈیٹا کی اہمیت: ڈیٹا لوگوں اور تنظیموں کو بہتر فیصلے کرنے میں مدد دینے میں بڑا کردار ادا کرتا ہے۔ ڈیٹا کے بغیر، فیصلے قیاس آرائیوں یا ذاتی رائے پر مبنی ہو سکتے ہیں جو غلطیوں کا باعث بن سکتے ہیں۔

مثال: اگر سکول انتظامیہ کوئی نیا اسپورٹس کلب شروع کرنا چاہتی ہے، تو وہ پہلے اس بارے میں ڈیٹا اکٹھا کر سکتے ہیں کہ طلبہ کو کون سے کھیل سب سے زیادہ پسند ہیں۔ اس سے انہیں ایک بہتر اور منصفانہ انتخاب کرنے میں مدد ملتی ہے۔

سوال 2: ڈیٹا سائنس کی حقیقی زندگی میں موجود اپیلی کیشنز (استعمالات) کی وضاحت کریں۔

جواب: ڈیٹا سائنس کی حقیقی زندگی میں موجود اپیلی کیشنز:

سوشل میڈیا: Instagram اور YouTube، Facebook جیسی اپیلی کیشنز ہمیں وہ پوسٹس اور ویڈیوز دکھانے کے لیے ڈیٹا استعمال کرتی ہیں جو ہمیں پسند آ سکتی ہیں۔

آن لائن شاپنگ: ای کامرس ویب سائٹس ہماری پچھلی خریداریوں کے ڈیٹا کا استعمال کرتے ہوئے مصنوعات تجویز کرتی ہیں۔
صحت کی دیکھ بھال (Healthcare): ہسپتال ڈیٹا سائنس کا استعمال بیماریوں کی پیش گوئی کرنے اور دوا کی ضرورت کی دیکھ بھال کو زیادہ مؤثر طریقے سے منظم کرنے کے لیے کرتے ہیں۔

نقل و حمل: Maps Google جیسی اپیلی کیشنز ٹریفک ڈیٹا کا استعمال کرتے ہوئے تیز ترین راستے تجویز کرتی ہیں۔
تعلیم: سکولوں میں، طلبہ کی حاضری اور امتحانی نتائج کے ڈیٹا کا استعمال سیکھنے کی حکمت عملیوں کو بہتر بنانے اور پیش رفت کو ٹریک کرنے کے لیے کیا جاتا ہے۔

ڈیٹا سائنس لائف سائیکل کا ایک جائزہ

5.2

سوال 3: ڈیٹا سائنس لائف سائیکل (DSL) کی تعریف کریں اور اس کے مراحل کے نام تحریر کریں۔

جواب: ڈیٹا سائنس ایک ہی مرحلے میں مکمل نہیں ہوتی۔ یہ مراحل کا ایک سلسلہ ہے جو کسی مسئلے کو حل کرنے میں مدد دیتا ہے۔ اسی کو ہم ڈیٹا سائنس لائف سائیکل (DSL) کہتے ہیں۔ DSL عام طور پر مندرجہ ذیل مراحل پر مشتمل ہوتا ہے:

مرحلہ 1 مسئلہ کو سمجھنا

مرحلہ 2 ڈیٹا اکٹھا کرنا

مرحلہ 3 ڈیٹا کی صفائی اور توثیق (Cleaning and Validating data)

مرحلہ 4 ڈیٹا کا تجزیہ کرنا

مرحلہ 5 ڈیٹا کی تشریح کرنا

مرحلہ 6 نتائج کو تصور (Visualize) کرنا اور مواصلت کرنا

سوال نمبر 4: ڈیٹا سائنس لائف سائیکل DSLC میں مسئلہ کو سمجھنا کی وضاحت کریں؟

جواب: ڈیٹا سائنس لائف سائیکل میں مسئلہ کو سمجھنا سب سے ابتدائی مرحلہ ہے۔ اس سے پہلے کہ ہم ڈیٹا اکٹھا کرنے کے بارے میں بات کرنا شروع کریں، زیر غور مسئلہ کو واضح طور پر سمجھنا ضروری ہے۔ اگر ہم مسئلہ کو صحیح طور پر نہیں سمجھیں گے تو ہم غلط ڈیٹا اکٹھا کر سکتے ہیں جس کے نتیجے میں بیکار یا غلط نتائج برآمد ہوں گے۔ اس ابتدائی مرحلے میں ہمیں ایسے سوالات پوچھنے کی ضرورت ہے جیسے:

- ہم کیا جاننے کی کوشش کر رہے ہیں؟
- یہ مسئلہ کیوں اہم ہے؟
- نتائج کون استعمال کرے گا؟

مسئلہ کو سمجھنے کے ذیلی حصے:

مسئلے کی شناخت اور تعریف: سب سے پہلے ہمیں اس مسئلے کی شناخت کرنی ہے جسے ہم ڈیٹا کے ذریعے حل کرنا چاہتے ہیں۔ مسئلہ وہاں متعلق ہو سکتا ہے جہاں معلومات کی بنیاد پر فیصلے کیے جاتے ہیں۔ ایک بار مسئلے کی نشاندہی ہو جانے کے بعد ہمیں اسے واضح اور سادہ انداز میں بیان کرنا چاہیے۔ مسئلے کی تعریف کرتے وقت ہمیں درج ذیل باتوں پر غور کرنے کی ضرورت ہے:

- (i) واضح اور سادہ الفاظ استعمال کریں
- (ii) ایک مخصوص سوال پر توجہ مرکوز کریں
- (iii) غیر واضح یا عمومی اصطلاحات سے گریز کریں۔

کیا آپ جانتے ہیں؟

ایک اچھا مسئلہ ہوتا ہے جس کا جواب ڈیٹا اکٹھا کر کے اور تجزیہ کر کے دیا جاسکے۔

واضح مقصد (Objective) کا تعین کرنا:

ایک بار جب ہم مسئلے کی نشاندہی اور تعریف کر لیں، تو ہمیں یہ طے کرنا ہوگا کہ ہم کیا حاصل کرنا چاہتے ہیں۔ انہیں ہمارے مقاصد کہتے ہیں، جو ڈیٹا سائنس کے باقی عمل کی رہنمائی میں مدد دیتے ہیں۔ وہ بتاتے ہیں:

- (i) ہمیں کس قسم کے ڈیٹا کی ضرورت ہے
- (ii) ہم کون سے سوالات پوچھیں گے
- (iii) ہم کون سے نتائج تلاش کر رہے ہیں۔

کیا آپ جانتے ہیں؟

اپنے مقاصد کو ایک چیک لسٹ کی طرح لکھیں۔ اس سے آپ کے پروجیکٹ پر توجہ مرکوز رکھنے میں مدد ملتی ہے۔

مثالیں: اسکول، صحت، کاروبار سے متعلق مسائل اور مقاصد

یہ ذیلی حصہ تصور کو بہتر طریقے سے سمجھنے کے لیے چند حقیقت پسندانہ مثالیں فراہم کرتا ہے۔ جدول 5.1 مختلف شعبوں جیسے اسکول، صحت اور کاروبار سے متعلق مثالی مسائل اور مقاصد فراہم کرتا ہے۔

جدول 5.1: مثالی مسائل اور مقاصد

ڈومین	مثالی مسئلہ	مقاصد
اسکول	بہت سے طالب علم اسکول دیر سے پہنچتے ہیں۔	ایک مینیجمنٹ سسٹم کے ذریعے طالب علم دیر سے پہنچتے ہیں، اس کا پتہ لگانا۔
		دیر سے آنے والی وجوہات کی شناخت کرنا (مثلاً، ٹرانسپورٹ، فاصلہ، ہنگامی صورتحال، یا موسمی مسائل)۔

صحت	سکول کے اوقات کے دوران بہت سے طالب علم کافی زیادہ پانی نہیں پیتے ہیں۔	طالب علم روزانہ کتنا پانی پیتے ہیں، اس پر ڈیٹا اکٹھا کرنا	دیر سے آنے والے کیمز کو کم کرنے کے لیے حل تجویز کرنا۔
کاروبار	سکول کی کینٹین یہ جاننا چاہتی ہے کہ طالب علم کون سے سٹیکس سب سے زیادہ پسند کرتے ہیں۔	طالب علموں کے پسندیدہ اسٹیکس کے بارے میں سروے کرنا۔	پانی کی کیم مقدار پر پانی کی ڈیٹا اکٹھا کرنا۔
		مقبولیت اور فروخت کے ڈیٹا کا موازنہ کرنا۔	صحت مند عادات کو فروغ دینے کے لیے اقدامات کی سفارش کرنا۔
		کینٹین مینیجر کو سمارٹ آرڈرنگ کے فیصلے کرنے میں مدد فراہم کرنا۔	

یاد رکھیں:

- 1- کسی بھی ڈیٹا سائنس پروجیکٹ کا پہلا مرحلہ مسئلہ کو سمجھنا ہوتا ہے۔
- 2- ہمیں مسئلہ کو واضح طور پر بیان کرنا اور مخصوص اہداف طے کرنے چاہئیں۔
- 3- کچھ مسائل وہ ہوتے ہیں جو ڈیٹا کے ذریعے حل کیے جاسکتے ہیں۔
- 4- واضح مقاصد کو ڈیٹا اکٹھا کرنے اور مفید نتائج تک پہنچنے میں مدد دیتے ہیں۔

سوال 5: ڈیٹا اکٹھا کرنے (Data Collection) کی تعریف کریں اور ڈیٹا اکٹھا کرنے کے طریقوں کی بھی وضاحت کریں؟
جواب: ڈیٹا خام مال کی طرح ہے، اور ہمیں اسے صاف کرنے، ذخیرہ کرنے یا تجزیہ کرنے کے لیے پہلے صحیح قسم کا ڈیٹا اکٹھا کرنا ہوگا۔ ڈیٹا اکٹھا کرنے کے لیے، ہم درج ذیل مراحل استعمال کرتے ہیں:

- ڈیٹا اکٹھا کرنے کے طریقے
- صحیح ماخذ کا انتخاب
- ڈیٹا کے معیار (Quality) اور تعصب (Bias) کا جائزہ
- ڈیٹا اکٹھا کرنے کے طریقے:
- مسئلے اور دستیاب آلات کے لحاظ سے ڈیٹا اکٹھا کرنے کے بہت سے طریقے ہیں۔
- سروے (آن لائن / آف لائن)۔

سروے سب سے آسان اور عام طریقہ ہے۔ ہم لوگوں سے سوالات پوچھتے ہیں اور ان کے جوابات ریکارڈ کرتے ہیں۔
مثال: یہ جاننے کے لیے کہ طلبہ کو کون سے کھیل پسند ہیں، ہم Google Form کا استعمال کرتے ہوئے سوالات پر مشتمل ایک سروے بنا سکتے ہیں اور اس کا لنک طلبہ کو جوابات کے لیے بھیج سکتے ہیں۔

کیا آپ جانتے ہیں؟
سروے گوگل فارمز (آن لائن) اور پرسنل شدہ کاغذی فارمز (میسکول) کا استعمال کرتے ہوئے کیے جاسکتے ہیں۔
APIs اور آن لائن ذرائع: آپلیکیشن پروگرام انٹرفیس (API) سافٹ ویئر کو ویب سائٹس یا ایپلیکیشنز (مختصر ایپس) سے خود کار طریقے سے ڈیٹا اکٹھا کرنے کی اجازت دیتے ہیں۔
مثال: ایک اسکول روزانہ حاضری کے ڈیٹا کو رپورٹ میں لانے کے لیے حاضری API استعمال کر سکتا ہے۔

IoT اور سنسر ڈیٹا: انٹرنیٹ آف تھنگز (IoT) ڈیوائسز جیسے سمارٹ واچز، پانی کے ڈسنسر، یا درجہ حرارت کے سنسر خود کار طور پر ڈیٹا اکٹھا کر سکتے ہیں۔
مثال: پانی پینے کی عادات کا مطالعہ کرنے کے لیے، اسکول قواعدوں کے قریب سمارٹ واٹر میٹر لگا سکتے ہیں۔

سمارٹ نیس ۱۰

IoT ڈیوائسز صحت، حفاظت اور ماحولیاتی نگرانی میں مفید ہیں۔

ویب اسکرپنگ (Web Scraping)

یہ ویب سائٹس سے عوامی ڈیٹا کو پروگرام کے ذریعے اکٹھا کرنے کا عمل ہے۔ مثال کے طور پر، ایک کاروباری طالب علم آن لائن اسٹورز سے اسٹیکس کی قیمتوں کو اسکرپ یا نکال سکتا ہے تاکہ ان کا موازنہ اسکول کی کیفینین کی قیمتوں سے کیا جاسکے۔

سمارٹ نیس ۱۰

ویب اسکرپنگ صرف ان ویب سائٹس کے لیے استاد کی اجازت سے کی جانی چاہیے جو اس کی اجازت دیتی ہوں۔

سوال 6: درست ڈیٹا سورس کا انتخاب اور ڈیٹا کے معیار اور تعصب کا جائزہ لینے کی وضاحت کریں؟

جواب: درست ڈیٹا سورس کا انتخاب:

تمام ڈیٹا مفید نہیں ہوتا۔ ہمیں قابل اعتماد، درست اور متعلقہ ذرائع کا انتخاب کرنا چاہیے۔ ڈیٹا ماخذ کا انتخاب کرتے وقت، چند سوالات پوچھیں جیسے (i) کیا یہ ڈیٹا کسی قابل اعتماد جگہ سے ہے؟ (ii) کیا یہ حالیہ ہے یا پرانا؟ اور (iii) کیا یہ اس مسئلے سے مطابقت رکھتا ہے جسے میں حل کرنا چاہتا ہوں؟
مثال: اگر ہم 2025 میں طلبہ کے پسندیدہ اسٹیکس جاننا چاہتے ہیں، تو ہمیں 2020 میں کیے گئے سروے کو استعمال کرنے سے گریز کرنا چاہیے۔
ڈیٹا کے معیار اور جانبداری کا جائزہ:

ڈیٹا کا معیار (Data Quality):

ڈیٹا کے معیار کا مطلب ہے کہ ڈیٹا:

(i) مکمل ہو (کوئی گم شدہ قدریں نہ ہوں)

(ii) درست ہو (کوئی غلطیاں نہ ہوں)

(iii) متعلقہ ہو (ہمارے مسئلے سے مطابقت رکھتا ہو)۔

جانبداری (Bias): جانبداری کا مطلب ہے کہ ڈیٹا سچائی کا صرف ایک حصہ دکھاتا ہے اور پوری تصویر سے محروم رہ جاتا ہے۔

مثال: اگر ہم صرف لڑکوں سے ان کے پسندیدہ کھیلوں کے بارے میں سروے کرتے ہیں، تو ہم لڑکیوں کی رائے سے محروم رہ جائیں گے۔ نتائج غیر منصفانہ یا لڑکوں کی طرف متعصبانہ ہوں گے۔

جانبداری سے بچنے کے اقدامات:

جانبداری سے بچنے کے لیے درج ذیل باتوں کو یاد رکھیں:

- ممنوع کردہ ذرائع سے پوچھیں۔
- سوالات کو غیر جانبدار رکھیں اور سرعوب سن نہ ہوں۔
- عجیب و غریب نمونوں کے لیے جوابات کا جائزہ لیں۔

سمارٹ نیس ۱۰

ہمیشہ تازہ اور مسئلے سے متعلق ڈیٹا اکٹھا کریں۔

ایک سادہ سروے ڈیزائن کرنا
مقصد: اپنے ہم جماعتوں کو فارغ وقت کی عادات کے بارے میں ایک مختصر سروے بنا کر ڈیٹا اکٹھا کرنے اور ذخیرہ کرنے کا طریقہ سیکھیں۔
ہدایات:

- 1- اپنا سروے ڈیزائن کریں: اپنے ہم جماعتوں سے پوچھنے کے لیے 3 سے 5 سوالات سوچیں۔
- 2- جوابات اکٹھا کرنے کی تیاری کریں: ایک طریقہ کا انتخاب کریں (مثلاً، کاغذی فارم یا گوگل فارم)۔
- 3- ڈیٹا اکٹھا کریں: کم از کم 5 ہم جماعتوں سے اپنا سروے مکمل کرنے کو کہیں اور ان کے جوابات کو واضح طور پر لکھیں۔
- 4- ڈیٹا کو ذخیرہ کریں: (i) اگر آپ کاغذ استعمال کر رہے ہیں، تو اسے صاف ستھرا اور کالموں میں منظم رکھیں۔ (ii) اگر ڈیجیٹل ٹولز استعمال کر رہے ہیں، تو فائل کو ایک واضح نام جیسے FreeTime-your name.xlsx سے محفوظ کریں۔
- 5- (اختیاری طور پر) اپنے استاد یا ہم جماعتوں کو دکھائیں کہ آپ نے ڈیٹا کیسے اکٹھا اور ذخیرہ کیا۔

یاد رکھیں!

- 1- مسئلے کی تعریف کرنے کے بعد، ہمیں وہ ڈیٹا اکٹھا کرنے کی ضرورت ہے جو مسئلے کو حل کرنے میں مدد دے سکے۔
- 2- ڈیٹا اکٹھا کرنے کے مختلف طریقے ہیں جیسے سروے، APIs، سینسرز، اور ویب اسکرپنگ۔
- 3- اگلے مرحلے پر جاننے سے پہلے ہمیں ڈیٹا کے معیار اور (تعبص) کی جانچ کرنی چاہیے۔
- 4- نامکمل یا متعصبانہ ڈیٹا غلط فیصلوں کا باعث بن سکتا ہے۔

سوال 7: ڈیٹا کلیننگ اور تصدیق (Cleaning and Validating Data) کی وضاحت کریں؟

جواب: ڈیٹا اکٹھا کرنے کے بعد، اگلا اہم مرحلہ اسے صاف کرنا اور توثیق کرنا ہے۔ یاد رکھیں کہ ہم نے پہلے سروے کی اور دیگر طریقوں سے ڈیٹا اکٹھا کیا تھا۔ اب، ہمیں یہ جانچنے کی ضرورت ہے کہ آیا وہ ڈیٹا درست، مکمل اور استعمال کے لیے تیار ہے۔ یہ مرحلہ اس بات کو یقینی بناتا ہے کہ ہم بعد میں جس ڈیٹا کا تجزیہ کریں گے وہ قابل اعتبار ہو۔

خام ڈیٹا (Raw Data): خام ڈیٹا وہ اصل معلومات ہے جو ہم لوگوں، سینسرز، ویب سائٹس یا ایپس سے اکٹھا کرتے ہیں۔ اسے ابھی چیک یا صاف نہیں کیا گیا ہو۔

مثال: فرض کریں کہ 20 طلبہ نے پسندیدہ کھیلوں کے بارے میں آپ کا سروے بھرا۔ کچھ اندراج غلط ہیں یا غلط لکھے گئے ہیں۔ خام ڈیٹا میں عام مسائل

سیریل نمبر	غلطی کی قسم	مثال
1	مینگ ڈیٹا	خالی خانے (مثلاً، کھیل کا نام یا وقت درج نہیں)۔
2	ڈوپلیکیٹ اندراجات	ایک ہی طالب علم کو دو بار وہ درج کیا جاتا۔
3	غیر مستحکم فارمیٹس (Inconsistent formats)	"Cricket"، "cricket"، "CRICKET" (حروف کی یکجۂ سازی میں فرق)۔
4	ججوں کی غلطیاں (Spelling mistakes)	"Football" کی بجائے "Footbal" یا "Daily" کی بجائے "Daliy"۔
5	غیر معمولی ویلیوز (Outliers) (عجیب و غریب اقدار)	فریکوئنسی = 1000 بار فی ہفتہ۔

- ڈیٹا تو شیق کا مطلب ہے یہ جانچنا کہ ڈیٹا:
- صحیح اصولوں کی پیروی کرتا ہے (مثلاً، صرف درست گھنٹوں کے نام)
 - قیمتوں کی صحیح قسم رکھتا ہے (مثلاً، جہاں اعداد کی توقع ہو وہاں اعداد ہوں)
 - جعلی یا بے وقوفانہ نہ ہو۔

تو شیق کی مثال (Example Validation Rule)

'پسندیدہ کھیل' کے لیے صرف ان میں سے کسی ایک قدر کی اجازت دیں:

"Favorite Sport": Football, Cricket, Badminton, Hockey

ڈیٹا کلیننگ کے طریقے (Data Cleaning Methods)

ایک بار جب ہمیں غلطیاں مل جائیں، تو ہمیں انہیں ٹھیک کرنے یا ہٹانے کی ضرورت ہے۔ اسے ڈیٹا صفائی کہتے ہیں۔
ڈیٹا صاف کرنے کے لیے درج ذیل چند بنیادی اقدامات ہیں:

- جیسے کی غلطیاں ٹھیک کریں: "Football" → "Footbell"
- فارمیٹس کو معیاری بنائیں: "Cricket" → "CRICKET"
- نقل شدہ اعداد و احوال کو ہٹائیں: ایک ہی طالب علم کے لیے بار بار کیے گئے اندراجات کو حذف کریں۔
- گرم شدہ قدریں چیلنج کریں: اگر تعداد 100 سے زیادہ ہو تو پوری قطار ہٹا دیں۔
- غلط ڈیٹا کو ٹھیک کریں: اگر تعداد 100 سے زیادہ ہو تو اسے ٹھیک کریں یا ہٹا دیں۔

سوال نمبر 8: ڈیٹا کا تجزیہ کرنا (Analyzing Data) کی وضاحت کریں؟ تجزیہ کے لیے آسان اوزار بیان کریں؟

جواب: اب جب کہ ہمارا ڈیٹا صحیح طریقے سے اکٹھا اور صاف کیا جا چکا ہے، ہم تجزیہ کے مراحل کی طرف بڑھ سکتے ہیں۔ یہ وہ مرحلہ ہے جہاں ہم ڈیٹا کو دیکھ کر پیٹرن، رجحانات، اور سوالات کے جوابات تلاش کرتے ہیں۔ ڈیٹا کا تجزیہ کرنے کا مطلب ہے ڈیٹا کو احتیاط سے جانچنا تاکہ (i) اس کا خلاصہ کیا جاسکے، (ii) مختلف گروہوں کا موازنہ کیا جاسکے، (iii) پیٹرن تلاش کیے جاسکیں۔
تجزیے کے لیے سادہ ٹولز کا استعمال:

آپ Google Sheets، Microsoft Excel یا نوٹ بک میں دستی تحریر جیسے سادہ ٹولز سے ڈیٹا کا تجزیہ شروع کر سکتے ہیں۔

مثال: آپ نے 20 طلبہ کا سروے کیا۔ جدول 5.4 دکھاتی ہے کہ آپ کو کیا ملا ہو سکتا ہے۔

پسندیدہ کھیل	فریکوئنسی
کرکٹ	10
فٹ بال	5
ہاکی	4
بلیڈمنٹن	1

اس ڈیٹا سے، ہم مشاہدہ کر سکتے ہیں کہ زیادہ تر طلبہ کرکٹ کو ترجیح دیتے ہیں۔

یاد رکھیں!

- 1- ڈیٹا تجزیہ ہمیں پیٹرن تلاش کرنے اور فیصلے کرنے میں مدد دیتا ہے۔ 2- ہم ڈیٹا کو سمجھنے کے لیے الگ الگ اوزار استعمال کرتے ہیں۔
- 3- یہاں تک کہ میزیں (ٹیبلز) یا اسپریڈ شیٹس جیسے آسان اوزار بھی ہمیں ڈیٹا کا تجزیہ کرنے میں مدد دے سکتے ہیں۔

سوال 9: ڈیٹا کی اہم پریشانی (Data Interpreting) کی وضاحت کریں؟ نیز ڈیٹا ویژولائزیشن (Data Visualization) اور نتائج کی مواصلت (Communication Results) کی وضاحت کریں؟

جواب: یہ مرحلہ ڈیٹا سے معنی تلاش کرنے (یعنی رجحانات اور پیٹرن کی نشاندہی کرنے کے بارے میں ہے۔ ڈیٹا کا خلاصہ کرنے کے بعد، ہم خود سے پوچھ سکتے ہیں:

- اعداد ہمیں کیا بتاتے ہیں؟
- کچھ آپشنز زیادہ مقبول کیوں ہیں؟
- کیا کوئی حیرت یا پیٹرن ہیں؟

مثال: اگر ہمیں پتہ چلے کہ زیادہ تر دیر سے آنے والے طلبہ اسکول سے دور رہتے ہیں، تو ہم بطور حل نقل و حمل کی سہولت فراہم کرنے کی سفارش کر سکتے ہیں۔ نتائج کو ویژولائز کرنا اور مواصلت کرنا (Visualizing and Communicating Results) لوگ عام طور پر اعداد سے بہتر تصاویر کو سمجھتے ہیں، اسی لیے ہم ڈیٹا تصورات کا استعمال کرتے ہیں۔

یاد رکھیں کہ پچھلے مراحل میں، ہم نے پایا تھا کہ کرکٹ سب سے مقبول کھیل تھا۔ اس مرحلے میں، ہم ایک چارٹ بنائیں گے جو یہ واضح طور پر دکھاتا ہے۔ اس سے ہم جماعتوں، اساتذہ یا دیگر افراد کو اپنے نتائج پیش کرنا آسان ہو جاتا ہے۔

ڈیٹا ویژولائزیشن کی اہمیت:

- ڈیٹا کو سمجھنے میں آسان بناتا ہے۔
- ایک نظر میں پیٹرن دکھاتا ہے۔
- پریزنٹیشنز یا رپورٹس میں خیالات کو واضح طور پر پیش کرنے میں مدد دیتا ہے۔

مثال: یہ کہنے کے بجائے کہ 10 طلبہ کرکٹ پسند کرتے ہیں، ہم اسے بار چارٹ میں دکھا سکتے ہیں۔

چارٹس اور گرافس کی عام اقسام (بار، پائی، لائن وغیرہ)

ہمیشہ اپنے ڈیٹا سے مماثل چارٹ کا انتخاب کریں۔ جدول 5.6 عام طور پر استعمال ہونے والی چارٹ کی اقسام، ان کے استعمال کی صورت حال اور مثالیں دکھاتی ہے۔

جدول 5.6: چارٹ کی اقسام اور استعمالات

چارٹ / گراف کی قسم	استعمال کا مقصد	مثال
بار چارٹ	گروپس کا موازنہ کرنا	ہر کھیل کو پسند کرنے والے طالب علموں کی تعداد۔
پائی چارٹ	پوری چیز کے حصے دکھانا	پسندیدہ اسٹیکس کا فیصد۔
لائن چارٹ	وقت کے ساتھ تبدیلی دکھانا	ہر ہفتے دیر سے آنے والے طالب علموں کی تعداد۔

سمارت ٹیپس:

اپنے چارٹس کو سادہ، صاف اور مناسب طور پر لیبل شدہ رکھیں۔

کیونیکٹنگ رزلٹ (Communicating Results)

چارٹ تیار ہونے کے بعد، اپنے نتائج کی وضاحت کریں۔ آپ 2-3 سطریں لکھ سکتے ہیں:

”زیادہ تر طلبہ (10) کرکٹ پسند کرتے ہیں، اس کے بعد فٹ بال (6) ہے۔ بہت کم طلبہ (1) بیڈمنٹن پسند کرتے ہیں۔“

سماٹ نہیں! ڈیٹا کی کہانی سنانے کے لیے بصری معادلوں اور الفاظ کو ایک ساتھ استعمال کریں (واضح، مختصر اور مفید)۔

یاد رکھیں!

- 1- ڈیٹا کو تصور کرنے کا مطلب ہے نتائج دکھانے کے لیے چارٹس یا گرافس کا استعمال۔
- 2- یہ ڈیٹا کو سمجھنے اور شیئر کرنے میں آسان بناتا ہے۔
- 3- اپنے ڈیٹا کی قسم کے لیے صحیح چارٹ استعمال کریں۔
- 4- اپنے بصری مواد کو ہیٹ لیبل اور ایکسلین کریں۔

ویڈیولائزیشن کے لیے ٹولز

5.3

سوال 10: ویڈیولائزیشن (Visualization) کے لیے ٹولز بیان کریں؟ چارٹ بنانے کا طریقہ کار تحریر کریں؟

جواب: ایک بار جب آپ جان لیں کہ کس قسم کا چارٹ استعمال کرنا ہے، اگلا مرحلہ اسے ٹولز کا استعمال کرتے ہوئے تخلیق کرنا ہے۔ اس مرحلے میں، ہم چارٹس تخلیق کرتے ہیں جس سے ڈیٹا کو سمجھنا اور پیش کرنا آسان ہو جاتا ہے۔ مخصوص سافٹ ویئر کا استعمال کرتے ہوئے چارٹس اور گرافس بنانا آسان ہے۔

Google Sheets اور Microsoft Excel میزوں (ٹیبلز) سے چارٹس اور گرافس بنانے کے لیے سادہ مگر طاقتور ٹولز ہیں۔

Google Sheets میں چارٹ بنانے کے اقدامات (Chart Steps to Create)

- اپنا ڈیٹا قطاروں اور کالموں میں درج کریں۔
- ڈیٹا کی رینج (حد) کو منتخب کریں۔
- "Insert" مینیو پر کلک کریں پھر "Chart" کو منتخب کریں۔
- چارٹ کی قسم (بار، پائی، لائن) منتخب کریں۔
- چارٹ کو ٹائٹل، لیبل اور رنگ کے ساتھ حسب ضرورت بنائیں (Customize)۔

اسپریڈ شیٹ سافٹ ویئر کے فوائد:

ہم اسپریڈ شیٹس کا استعمال اس لیے کرتے ہیں کیونکہ وہ:

- استعمال میں آسان ہیں
- Google Sheets کے لیے (تنصیب (installation) کی ضرورت نہیں
- چھوٹے سے درمیانے درجے کے پروجیکٹس کے لیے اچھی ہیں۔

Frequency or No. of Students



فصل 5.1: طلبہ کے پسندیدہ کھیل (جدول 5.5 کے ڈیٹا کی بنیاد پر)

1- Google Sheets اور Excel جاننے کے لیے آسان ہوزار ہیں۔

2- ایک ڈیش بورڈ واضح ڈیٹا کہانی بتانے کے لیے ایک جگہ متعدد تصویری مواد دکھاتا ہے۔

حقیقی زندگی میں DSLC کا استعمال (Use of DSLC in real life)

معیاری DSLC چھ اہم مراحل پر مشتمل ہوتا ہے، تاہم، حقیقی دنیا کی کچھ اپیلی کیشنز/منظر ناموں میں، کچھ اضافی مراحل شامل کیے جاسکتے ہیں۔ ایسا ہی ایک مرحلہ ڈیٹا کو ذخیرہ کرنا اور منظم کرنا (Storing and Organizing Data) ہو سکتا ہے، خاص طور پر بڑے پروجیکٹس یا ٹیم سٹریکچر میں جہاں ڈیٹا بیسز اور منظم ذخیرہ اہم ہوتے ہیں۔ یہ عام طور پر DBMS اور SQL جیسے نظاموں کے ذریعے منظم کیے جاتے ہیں۔

5.4 ڈیٹا بیس اور ڈیٹا کو محفوظ کرنا

سوال 11: ڈیٹا بیس (Database) کی تعریف کریں؟ ہمیں سٹرکچرڈ ڈیٹا (structured data) کی ضرورت کیوں ہے؟ اسپریڈ شیٹ اور ڈیٹا بیس کے درمیان فرق بتائیں؟

جواب: ڈیٹا بیس ڈیٹا کا ایک منظم ذخیرہ (organized collection) ہوتا ہے۔ یہ ڈیٹا کو ٹیبلوں کی صورت میں ذخیرہ کرتا ہے۔ ہر ٹیبل ایک چھوٹی اسپریڈ شیٹ کی طرح ہوتی ہے جس میں قطاریں (rows) اور کالم (columns) ہوتے ہیں۔ ڈیٹا بیس ڈیٹا کو مؤثر طریقے سے ذخیرہ کرنے، حاصل کرنے (retrieve) اور اپ ڈیٹ کرنے میں مدد دیتی ہیں۔

ڈیٹا بیس کا مقصد (Purpose of Database)

ایک بے ترتیب سروے ٹیبل کو صاف کرنے کے بعد، صاف شدہ ڈیٹا کو مناسب فارمیٹ میں (یا تو اسپریڈ شیٹ میں یا ڈیٹا بیس میں) ذخیرہ کرنے کی ضرورت ہوتی ہے، تاکہ یہ منظم، قابل تلاش (searchable) اور تجزیہ کے لیے تیار ہو۔ سٹرکچرڈ ڈیٹا کو ذخیرہ کرنے کے دو اہم طریقے ہیں: (i) اسپریڈ شیٹس (ii) ڈیٹا بیسز

ڈیٹا بیس اور اسپریڈ شیٹ کے درمیان فرق (Difference between Database and Spreadsheet)

طریقہ	وضاحت	مثالی ٹول
اسپریڈ شیٹس	چھوٹے ڈیٹا بیس کے لیے قطاروں اور کالموں پر مشتمل ایک گروڈ۔	ایکسل (Excel)، گوگل شیٹس (Google Sheets)
ڈیٹا بیسز	بڑے ڈیٹا کو ٹیبلوں میں محفوظ کرنے اور منظم کرنے کے لیے ایک سٹرکچرڈ نظام۔	مائی ایس کیو ایل (MySQL)، ایس کیو لائٹ (SQLite)

ڈیٹا بیس کی اصطلاحات: معنی اور مثالیں

اصطلاح	معنی	مثال
ٹیبل	متعلقہ ڈیٹا کا ایک گروپ۔	اسپورٹس سروے
ریکارڈ	ٹیبل میں ایک واحد اندراج (قطار - row)۔	علی کے ڈیٹا کی مکمل قطار۔
فیلڈ	معلومات کا ایک مخصوص حصہ (کالم - column)۔	نام، پتہ، پروفیشنل

سمارٹ ٹیبل

اسپریڈ شیٹس چھوٹے پراجیکٹس کے لیے اچھی ہیں۔ ڈیٹا بیسز بڑے یا ملٹی یوزر ڈیٹا سسٹمز (multi-user data systems) کے لیے بہتر ہوتے ہیں۔

سوال 12: DBMS کی تعریف کریں؟ نیز SQL میں لکھے گئے چند استفسارات (queries) کی وضاحت کریں؟
 جواب: ڈیٹا بیس منیجمنٹ سسٹم (DBMS) ایک سافٹ ویئر ہے جو ڈیٹا کو تخلیق کرنے، منظم کرنے اور ان کے ساتھ تعامل کرنے کے لیے استعمال ہوتا ہے۔ مثالوں میں MySQL، SQLite، اور Microsoft Access شامل ہیں۔
 ڈیٹا بیس میں ڈیٹا کے ساتھ کام کرنے کے لیے، ہم ایک زبان استعمال کرتے ہیں جسے Structured Query Language (SQL) کہتے ہیں۔ SQL سادہ کمانڈز کا استعمال کرتے ہوئے ڈیٹا شامل کرنے، اپ ڈیٹ کرنے، حذف کرنے یا حاصل کرنے کے لیے ہماری مدد کرتا ہے۔ کچھ بنیادی SQL استفسارات درج ذیل ہیں:

SELECT * FROM Sports

یہ سادہ استفسار Sports نامی ٹیبل سے تمام ڈیٹا دکھاتا ہے۔

SELECT Name FROM Sports WHERE Favorite_Sport = 'Cricket';

یہ استفسار Sports ٹیبل میں موجود طلبہ کے نام دکھاتا ہے جو کرکٹ پسند کرتے ہیں۔

INSERT INTO Sports (Name, Favorite_Sport) VALUES ('Ali', 'Football');

یہ استفسار ایک نیا اندراج (entry) شامل کرتا ہے۔

یہ بنیادی SQL شیٹیں بڑے ڈیٹا بیسز سے مخصوص ڈیٹا کو مؤثر طریقے سے منظم کرنے اور حاصل کرنے میں مدد کرتی ہیں۔

یاد رکھیں!

- 1- آسان رسائی اور تجزیہ کے لیے ڈیٹا کو منظم (structured) طریقے سے ذخیرہ کیا جاتا ہے۔
- 2- اسپریڈ شیٹس چھوٹے ڈیٹا بیس کے لیے مفید ہیں۔
- 3- ڈیٹا بیسز بڑے اور زیادہ پیچیدہ ڈیٹا کے لیے استعمال ہوتے ہیں۔
- 4- ایک ٹیبل ریکارڈز (قطاریں) اور فیلڈز (کالم) کا استعمال کرتے ہوئے ڈیٹا ذخیرہ کرتی ہے۔
- 5- منظم ڈیٹا منظمی (duplication) کی غلطیوں کو کم کرتا ہے اور تجزیہ کو آسان بناتا ہے۔

معروضی طرز سوالات

اضافی کثیر الانتخابی سوالات

محکمہ تعلیم کی نئی امتحانی تکنیکس (Knowledge, Understanding, Application, Analytical & Conceptual) کی روشنی میں مرتب کیے گئے کثیر الانتخابی سوالات

تعارف

- 1- ڈیٹا سائنس کا مقصد:
 - (A) ڈیٹا کو باضابطہ طریقے سے پیش کرنا
 - (B) ڈیٹا کو معلومات میں تبدیل کرنا
 - (C) دونوں اور ب
 - (D) ان میں سے کوئی نہیں
- 2- ڈیٹا اکٹھا کرنے کا ذریعہ کون سا ہے؟
 - (A) سروے
 - (B) سوشل میڈیا
 - (C) موبائل ایپس
 - (D) تمام بالا

3- ڈیٹا سائنس کا کون سا فائدہ نہیں ہے؟

- (A) موسم کی پیش گوئی (B) نقدی کا حساب لگانا (C) صحت کی خدمات (D) کنٹرول

4- آرگنائزیشن ڈیٹا سائنس کیوں استعمال کرتی ہیں؟

- (A) فیصلہ سازی (B) مصنوعات کی ترقی (C) مشینوں کو کنٹرول کرنا (D) ان میں سے کوئی نہیں

ڈیٹا سائنس کا تعارف

5.1

5- ڈیٹا سائنس کے لیے کون سا ڈیٹا مفید نہیں ہے؟

- (A) فٹنس ٹریکنگ ڈیٹا (B) گوگل سرچ (C) آن لائن شاپنگ (D) ان میں سے کوئی نہیں

6- ڈیٹا سائنس کی تکنیک کون سی ہے؟

- (A) ترتیب دینا (B) تجزیہ کرنا (C) تصویری شکل دینا (D) تمام بالا

7- کس چیز سے غلطیاں ہو سکتی ہیں؟

- (A) ڈیٹا کا تجزیہ (B) اندازہ لگانا (C) بھری شکل دینا (D) سروے کے ذریعے جمع کرنا

8- مصنوعات کی سفارش _____ پلیٹ فارم کرتا ہے:

- (A) آبل میڈیا (B) صحت کی دیکھ بھال (C) نقل و حمل (D) آن لائن شاپنگ

9- ڈیٹا سائنس کی کون سی ایپلی کیشن نہیں ہے؟

- (A) تعلیم (B) نقل و حمل (C) کرنسی نوٹوں کا جمع کرنا (D) صحت کی دیکھ بھال

10- ڈیٹا سائنس کی کون سی ایپلی کیشن سیکھنے میں بہتری کی پیش رفت کو ٹریک کر کے اسے استعمال ہوتی ہے؟

- (A) تعلیم (B) نقل و حمل (C) صحت کی دیکھ بھال (D) آن لائن شاپنگ

ڈیٹا سائنس لائف سائیکل کا ایک جائزہ

5.2

11- DSLC کا مطلب ہے:

- (A) ڈیٹا سائنس لائف سائیکل (B) ڈیٹا سائنس لارج سائیکل (C) ڈیٹا سائنس لائف سائیکل (D) ڈیٹا سائنس لائف سائیکل

12- DSLC میں مراحل کی کل تعداد:

- (A) 2 (B) 4 (C) 6 (D) 8

13- ڈیٹا سائنس لائف سائیکل کس چیز میں مددگار ہے؟

- (A) ڈیٹا اکٹھا کرنا (B) مسئلہ حل کرنا (C) سروے (D) تجزیہ

14- مسئلہ کو سمجھنے کا مطلب ہے:

- (A) ڈیٹا اکٹھا کرنا (B) ڈیٹا کا تجزیہ (C) مسئلے کی اہمیت (D) ان میں سے کوئی نہیں

15- مسئلہ حل کرنے کے لیے واضح ہدف بنانا کہلاتا ہے:

- (A) ڈیٹا کو سمجھنا (B) ڈیٹا اکٹھا کرنا (C) ڈیٹا کا تجزیہ کرنا (D) ڈیٹا کی تشریح کرنا

16- جب ہم مسئلے کی وضاحت کرتے ہیں، تو ہم _____

- (A) واضح الفاظ استعمال کرتے ہیں (B) ہدف پر توجہ مرکوز کرتے ہیں (C) غیر واضح سے بچتے ہیں (D) تمام بالا

- 17- واضح مقصد طے کرنے کی اہمیت:
 (A) لاگت کا حساب لگانا
 (B) ڈیٹا سائنس کے عمل کی رہنمائی کرنا
 (C) شیڈول دکھانا
 (D) اہم اہم کی وضاحت کرنا
- 18- کون سی چیز صارف کو پروجیکٹ پر توجہ مرکوز کرنے دیتی ہے؟
 (A) مسئلے کی پہچان
 (B) واضح مقصد طے کرنا
 (C) دونوں اورب
 (D) ان میں سے کوئی نہیں
- 19- مسئلے کی پہچان کے بعد اگلا مرحلہ:
 (A) ڈیٹا اکٹھا کرنا
 (B) ڈیٹا کلیننگ
 (C) ڈیٹا تجزیہ
 (D) ویزولائزیشن
- 20- وہ طریقہ جس میں ہم سوال پوچھتے ہیں اور جوابات ریکارڈ کرتے ہیں:
 (A) IoT اور سینسر ڈیٹا
 (B) APIs ماخذ
 (C) سروے
 (D) ویب سکرپٹنگ
- 21- آن لائن سروے کرنے کے لیے بہترین ٹول کون سا ہے؟
 (A) سوشل میڈیا
 (B) گوگل فارم
 (C) ایم ایس ورڈ
 (D) گوگل کویب
- 22- کون سی چیز سافٹ ویئر کو ویب سائٹس یا ایپلی کیشنز سے ڈیٹا اکٹھا کرنے دیتی ہے؟
 (A) سروے
 (B) IoT اور سینسر ڈیٹا
 (C) APIs
 (D) ویب سکرپٹنگ
- 23- IoT کس چیز میں مفید ہے؟
 (A) نقل و حمل
 (B) صحت اور حفاظت
 (C) تعلیم
 (D) آن لائن شاپنگ
- 24- ڈیٹا اکٹھا کرنے کا عوامی طریقہ ہے:
 (A) سروے
 (B) IoT اور سینسر
 (C) ویب سکرپٹنگ
 (D) APIs
- 25- ڈیٹا صرف سچائی کا کچھ حصہ دکھاتا ہے اور مکمل تصویر سے محروم رہتا ہے:
 (A) نامکمل
 (B) جانبداری
 (C) قابل بھروسہ
 (D) متعلقہ
- 26- خام ڈیٹا اکثر ہوتا ہے:
 (A) بے ترتیب
 (B) نامکمل
 (C) غیر متعلقہ
 (D) تمام بالا
- 27- ڈیٹا کی تصدیق کس چیز کے لیے اہم ہے؟
 (A) کلیننگ
 (B) ترتیب
 (C) تجزیہ
 (D) متعلقہ
- 28- وہ ڈیٹا جسے چیک یا صاف نہ کیا گیا ہو:
 (A) خام ڈیٹا
 (B) قابل بھروسہ ڈیٹا
 (C) متعلقہ ڈیٹا
 (D) مکمل ڈیٹا
- 29- عام ڈیٹا کی غلطیوں کی کتنی اقسام ہیں؟
 (A) 1
 (B) 2
 (C) 3
 (D) 4
- 30- ڈیٹا کی توثیق کی تکنیکوں کی کتنی اقسام ہیں؟
 (A) 1
 (B) 2
 (C) 3
 (D) 4
- 31- ٹول ڈیٹا کی توثیق کرنے میں مدد دیتا ہے:
 (A) ایم ایس ورڈ
 (B) ایم ایس ایکسل
 (C) گوگل شیٹ
 (D) B اور C دونوں

- 32- ریڈیٹسی بیٹنگ کس مرحلے میں کی جاتی ہے؟
 (A) ڈیٹا اکٹھا کرنا (B) ڈیٹا تجزیہ (C) ڈیٹا کلیئج (D) ڈیٹا ویڈیو لائزیشن
- 33- ڈیٹا صفائی کے بعد اگلا مرحلہ ہے:
 (A) ڈیٹا تجزیہ (B) ڈیٹا تشریح (C) ڈیٹا پیری شکل (D) ان میں سے کوئی نہیں
- 34- رجحانات اور پیٹرن تلاش کرنا کس مرحلے میں شامل ہے:
 (A) ڈیٹا تجزیہ (B) ڈیٹا تشریح (C) ویڈیو لائزیشن (D) کلیئج
- 35- تجزیے کے لیے کون سا ٹول استعمال ہوتا ہے؟
 (A) ایم ایس ورڈ (B) کورل ڈرا (C) گوگل شیٹ (D) گوگل فارم
- 36- ڈیٹا سے معنی تلاش کرنے کے لیے کون سا طریقہ استعمال ہوتا ہے؟
 (A) ڈیٹا تجزیہ (B) ڈیٹا تشریح (C) ڈیٹا ویڈیو لائزیشن (D) ان میں سے کوئی نہیں
- 37- DSLC طریقے کا آخری مرحلہ:
 (A) ڈیٹا ویڈیو لائزیشن (B) ڈیٹا تشریح (C) ڈیٹا تجزیہ (D) ان میں سے کوئی نہیں
- 38- نتائج کا اشتراک کرنا کہلاتا ہے:
 (A) تشریح (B) کیونیکیشن (C) پیش کش (D) وضاحت

5.3 ویڈیو لائزیشن کے لیے ٹولز

- 39- ڈیٹا ویڈیو لائزیشن دینے کا ٹول ہے:
 (A) کورل ڈرا (B) فوٹو شاپ (C) ایم ایس ایکسل (D) ایم ایس ورڈ
- 40- ڈیٹا تصویری شکل دینے کے اوزار کس چیز میں مدد کرتے ہیں؟
 (A) ڈیٹا تجزیہ (B) ڈیٹا پیش کش (C) فیصلہ سازی (D) تمام بالا
- 41- چارٹس اور گرافس میں ڈیٹا کس کے ذریعے داخل کیا جاتا ہے؟
 (A) ٹیبلو (B) تصاویر (C) پیرا گراف (D) رپورٹس
- 42- ہسٹوگرام ایک ہے:
 (A) گراف (B) چارٹ (C) A اور B دونوں (D) ان میں سے کوئی نہیں

5.4 ڈیٹا میں اور ڈیٹا کو محفوظ کرنا

- 43- ڈیٹا کو محفوظ اور منظم کرنا کہلاتا ہے:
 (A) ڈیٹا اکٹھا کرنا (B) ڈیٹا میں (C) ڈیٹا کلیئج (D) اسٹوریج
- 44- SQL ایک ہے:
 (A) سروے سافٹ ویئر (B) ڈیٹا کلیئج سافٹ ویئر (C) ڈیٹا میں (D) ان میں سے کوئی نہیں
- 45- ڈیٹا محفوظ کرنے کے لیے بہترین ٹول کون سا ہے؟
 (A) ایم ایس ایکسل (B) گوگل شیٹ (C) ڈیٹا میں (D) گوگل فارم

- 46- کون سا سختی نظام بڑا ڈیٹا محفوظ کرتا ہے؟
 (A) اسپریڈ شیٹ (B) مائیکروسافٹ (C) مائی ایس کیو ایل (D) تمام بالا
- 47- ڈیٹا میں ڈیٹا کو کس شکل میں محفوظ کرتا ہے؟
 (A) رپورٹس (B) ٹیبلز (C) تصاویر (D) پیراگراف
- 48- ایم ایس ایس ایکس ایک ہے:
 (A) ڈیٹا بیس (B) ڈیٹا سٹرکچر (C) ڈی بی ایم ایس (D) ایس کیو ایل
- 49- مندرجہ ذیل میں سے کون سا DBMS ہے؟
 (A) ایم ایس ایکسل (B) ایم ایس ورڈ (C) ایم ایس پاورپوائنٹ (D) مائی ایس کیو ایل
- 50- DBMS میں، قطاروں کو بھی کہتے ہیں:
 (A) فیلڈز (B) کالم (C) سیز (D) ریکارڈز

جوابات:

A -10	C -9	D -8	B -7	D -6	D -5	A -4	B -3	D -2	C -1
C -20	A -19	B -18	B -17	D -16	A -15	C -14	B -13	C -12	C -11
C -30	C -29	A -28	C -27	D -26	B -25	C -24	B -23	C -22	B -21
D -40	C -39	B -38	A -37	B -36	C -35	A -34	A -33	C -32	D -31
D -50	D -49	C -48	B -47	C -46	C -45	C -44	B -43	A -42	A -41

اضافی مختصر سوالات

میکر تعلیم کی نئی امتحانی تکلیفیں (Knowledge, Understanding, Application, Analytical & Conceptual) کی روشنی میں مرتب کیے گئے مختصر جوابی سوالات

تعارف

- 1- ڈیٹا سائنس کی تعریف کریں۔
 جواب: ڈیٹا سائنس مسائل کو سمجھنے، نمونوں کو دریافت کرنے اور بہتر فیصلے کرنے کے لیے ڈیٹا استعمال کرنے کا عمل ہے۔
 مثال: ڈاراز آن لائن شاپنگ ایپ اپنی فروخت بڑھانے کے لیے ڈیٹا سائنس استعمال کرتی ہے۔
- 2- ڈیٹا سائنس کیوں اہم ہے؟
 جواب: ڈیٹا سائنس اس تمام ڈیٹا کو مفید معلومات میں تبدیل کرنے میں مدد دیتی ہے۔ چاہے موسم کی پیش گوئی ہو، طلبہ کی دلچسپیوں کو سمجھنا ہو، یا صحت کی دیکھ بھال کو بہتر بنانا ہو، ڈیٹا سائنس ہمارے مینے دیکھنے اور کام کرنے کے طریقے کو بدل رہی ہے۔
- 3- کاروباری تنظیموں کے ذریعے ڈیٹا سائنس کی تکنیک کیسے استعمال ہوتی ہے؟
 جواب: پوری دنیا میں تنظیمیں بہتری کے لیے ڈیٹا استعمال کرتی ہیں:
 • فیصلہ سازی • اخراجات میں بچت • بہتر خدمات تخلیق کرنا

4- DSLC کے مراحل کے نام لکھیں۔

جواب: ڈیٹا سائنس لائف سائیکل میں شامل ہیں۔

- مسئلے کو سمجھنا
- ڈیٹا اکٹھا کرنا اور صفایا کرنا
- ڈیٹا کا تجزیہ کرنا
- نتائج کو ریپورٹ لائزیشن اور ان کا ابلاغ کرنا
- ڈیٹا کی وضاحت کرنا

ڈیٹا سائنس کا تعارف

5.1

5- ڈیٹا سائنس کا مقصد کیا ہے؟

جواب: ڈیٹا سائنس وہ شعبہ ہے جہاں ہم ڈیٹا کو پتھرین تلاش کرنے، سوالوں کے جواب دینے، اور لوگوں یا کاروباروں کو بہتر فیصلے کرنے میں مدد دینے کے لیے استعمال کرتے ہیں۔ اس میں ڈیٹا اکٹھا کرنا، منظم کرنا، تجزیہ کرنا اور بصری شکل دینا شامل ہے۔

6- فیصلہ سازی میں ڈیٹا کیوں اہم ہے؟

جواب: ڈیٹا لوگوں اور تنظیموں کو بہتر فیصلے کرنے میں مدد دینے میں بڑا کردار ادا کرتا ہے۔ ڈیٹا کے بغیر، فیصلے اندازوں یا ذاتی رائے پر مبنی ہو سکتے ہیں جو غلطیوں کا باعث بن سکتے ہیں۔

مثال: اگر اسکول انتظامیہ ایک نیا اسپورٹس کلب شروع کرنا چاہتی ہے، تو وہ پہلے اس بارے میں ڈیٹا اکٹھا کر سکتے ہیں کہ طلبہ کو کون سے کھیل سب سے زیادہ پسند ہیں۔ اس سے انہیں ایک سمجھدار اور منصفانہ انتخاب کرنے میں مدد ملتی ہے۔

7- سوشل میڈیا اور آن لائن شاپنگ میں ڈیٹا سائنس کی حقیقی زندگی میں استعمال کی وضاحت کریں۔

جواب: سوشل میڈیا:

Facebook، YouTube اور Instagram جیسی ایپلیکیشنز ہمیں وہ پوسٹس اور ویڈیوز دکھانے کے لیے ڈیٹا استعمال کرتی ہیں جو ہمیں پسند ہو سکتی ہیں۔

آن لائن شاپنگ: ای کامرس ویب سائٹس ہماری پچھلی خریداریوں کے ڈیٹا کا استعمال کرتے ہوئے مصنوعات تجویز کرتی ہیں۔

8- صحت سائنس اور تعلیم میں ڈیٹا سائنس کی حقیقی زندگی میں استعمال کی وضاحت کریں۔

جواب: صحت کی دیکھ بھال:

ہسپتال بیماریوں کی پیش گوئی کرنے اور مریضوں کی دیکھ بھال کو زیادہ مؤثر طریقے سے منظم کرنے کے لیے ڈیٹا سائنس استعمال کرتے ہیں۔
تعلیم: اسکولوں میں، طلبہ کی حاضری اور امتحان کے نتائج کا ڈیٹا سیکھنے کی حکمت عملیوں کو بہتر بنانے اور پروفٹ کوڑیک کرنے کے لیے استعمال کیا جاتا ہے۔

ڈیٹا سائنس لائف سائیکل کا ایک جائزہ

5.2

9- DSLC کی تعریف کریں اور اس کے مراحل کے نام لکھیں۔

جواب: ڈیٹا سائنس ایک ہی مرحلے میں مکمل نہیں ہوتی۔ یہ اقدامات کا ایک سلسلہ ہے جو کسی مسئلے کو حل کرنے میں مددگار ہوتے ہیں۔ اسے ہم ڈیٹا سائنس لائف سائیکل (DSLC) کہتے ہیں۔ DSLC عام طور پر مندرجہ ذیل مراحل پر مشتمل ہوتا ہے:

- مسئلے کو سمجھنا
- ڈیٹا اکٹھا کرنا
- ڈیٹا کی صفائی اور توثیق
- ڈیٹا کا تجزیہ کرنا
- ڈیٹا کی وضاحت کرنا
- نتائج کی ویپورٹ لائزیشن اور ان کا ابلاغ

10- مسئلے کو سمجھنا کیوں ضروری ہے؟

جواب: زیر غور مسئلے کو واضح طور پر سمجھنا ضروری ہے۔ اگر ہم مسئلے کو صحیح طریقے سے نہیں سمجھتے ہیں، تو ہم غلط ڈیٹا اکٹھا کر سکتے ہیں جس کے نتیجے میں بیکار یا غلط نتائج برآمد ہوں گے۔

11- مسئلے کی نشاندہی کرنے کے لیے کتنے مراحل شمار کیے جاتے ہیں؟

جواب: مسئلے کی تعریف کرتے وقت ہمیں درج ذیل چار عوامل کی ضرورت ہے:

(i) واضح اور آسان الفاظ استعمال کریں

(ii) ایک مخصوص سوال پر توجہ مرکوز کریں

(iii) غیر واضح یا عمومی اصطلاحات سے بچیں۔

12- واضح مقصد طے کرنے سے کیا مراد ہے؟

جواب: ایک بار جب ہم نے مسئلے کی نشاندہی اور تعریف کر لی، تو ہمیں فیصلہ کرنا ہوگا کہ ہم کیا حاصل کرنا چاہتے ہیں۔ انہیں ہمارے مقاصد کہا جاتا ہے جو ڈیٹا سائنس کے باقی عمل کی رہنمائی میں مدد کرتے ہیں۔ وہ بتاتے ہیں:

(i) ہمیں کس قسم کے ڈیٹا کی ضرورت ہے

(ii) ہم کون سے سوالات پوچھیں گے

(iii) ہم کون سے نتائج تلاش کر رہے ہیں۔

13- مسئلے کو واضح طور پر سمجھنے کے کیا فوائد ہیں؟

جواب:

- اچھی طرح طے شدہ مسائل ڈیٹا کا استعمال کرتے ہوئے حل کیے جاسکتے ہیں۔
- واضح مقاصد صحیح ڈیٹا اکٹھا کرنے اور مفید نتائج تک پہنچنے میں مدد کرتے ہیں۔

14- ڈیٹا اکٹھا کرنے کی تعریف کریں اور ڈیٹا اکٹھا کرنے کے طریقوں کے مراحل کے نام بتائیں۔

جواب: ڈیٹا خام مال کی طرح ہے، اور ہمیں اسے صاف کرنے، ذخیرہ کرنے یا تجزیہ کرنے سے پہلے صحیح قسم کا ڈیٹا اکٹھا کرنا ہوگا۔ ڈیٹا اکٹھا کرنے کے طریقے:

• سروے • APIs اور آن لائن ذرائع • IoT سسٹمز • ویب اسکرپنگ

15- مثال کے ساتھ سروے کی تعریف کریں۔

جواب: سروے سب سے آسان اور عام طریقہ ہے۔ ہم لوگوں سے سوالات پوچھتے ہیں اور ان کے جوابات ریکارڈ کر لیتے ہیں۔ مثال: یہ جاننے کے لیے کہ طلبہ کو کون سے کھیل پسند ہیں، ہم Google فارم کا استعمال کرتے ہوئے سوالات کے ساتھ ایک سروے بنا سکتے ہیں اور جوابات کے لیے اس کا لنک طلبہ کے ساتھ شیئر کر سکتے ہیں۔

16- APIs اور آن لائن ذرائع کا مقصد کیا ہے؟

جواب: ایپلیکیشن پروگرام انٹرفیس (APIs) سافٹ ویئر کو ویب سائنس یا ایپلیکیشنز سے خود بخود ڈیٹا اکٹھا کرنے کی اجازت دیتے ہیں۔ مثال: ایک اسکول روزانہ حاضری کے ڈیٹا کو ایک رپورٹ میں پل کرنے کے لیے حاضری API استعمال کر سکتا ہے۔

17- IoT اور سنسر ڈیٹا کی تعریف کریں۔

جواب: انٹرنیٹ آف تھنگز (IoT) ذیوائسز جیسے سمارٹ گھڑیاں، پانی دینے والے آلات، یا درجہ حرارت کے سنسر خود بخود ڈیٹا اکٹھا کر سکتے ہیں۔ مثال: پینے کی پانی کی عادات کا مطالعہ کرنے کے لیے، اسکول فواروں کے قریب سمارٹ واٹر میٹر لگا سکتے ہیں۔

18- ویب اسکرپنگ کیا ہے؟

جواب: یہ ویب سائنس سے پروگرام کی طور پر خواہی ڈیٹا اکٹھا کرنے کا ایک عمل ہے۔ مثال کے طور پر، ایک کاروباری طالب علم آن لائن اسٹورز سے اسٹیکس کی قیمتیں نکال سکتا ہے تاکہ ان کا اسکول کے بزنس کی قیمتوں سے موازنہ کیا جاسکے۔

19- ڈیٹا کو الٹی کیا ہے؟

جواب: ڈیٹا کو الٹی کا مطلب ہے کہ ڈیٹا:

(i) مکمل ہے (کوئی گم شدہ اقدار نہیں)

(ii) درست ہے (کوئی غلطی نہیں)

(iii) متعلقہ ہے (ہمارے مسئلے سے میل کھاتا ہے)۔

20- ڈیٹا اکٹھا کرنے میں جانبداری کیا ہے؟

جواب: جانبداری کا مطلب ہے کہ ڈیٹا سچائی کا صرف ایک حصہ دکھاتا ہے اور مکمل تصویر سے محروم رہ جاتا ہے۔

مثال: اگر ہم صرف لڑکوں سے ان کے پسندیدہ کھیلوں کے بارے میں سروے کرتے ہیں، تو ہم لڑکیوں کی رائے سے محروم رہ جائیں گے۔ نتائج غیر منصفانہ یا لڑکوں کی طرف جانبداری ہوں گے۔

21- جانبداری سے بچنے کے لیے کیا اقدامات کیے جاتے ہیں؟

جواب: جانبداری سے بچنے کے لیے درج ذیل یاد رکھیں:

- لوگوں کے متنوع گروپ سے پوچھیں۔
- عجیب پیرن کے لیے جوابات کا جائزہ لیں۔
- سوالات کو غیر جانبدار اور غیر راہنما رکھیں۔

22- خام ڈیٹا کیا ہے؟ ایک مثال دیں۔

جواب: خام ڈیٹا وہ اصل معلومات ہے جو ہم لوگوں، سینرز، ویب سائٹس یا ایپس سے اکٹھا کرتے ہیں۔ اسے ابھی چیک یا صاف نہیں کیا گیا ہے۔

مثال: فرض کریں کہ 20 طلبہ نے پسندیدہ کھیلوں کے بارے میں آپ کا سروے بھرا۔ کچھ اندراجات غائب ہیں یا غلط لکھے گئے ہیں۔

23- ڈیٹا تصدیق کی تکنیکوں کی وضاحت کریں۔

جواب: ڈیٹا تصدیق کا مطلب ہے یہ جانچنا کہ آیا ڈیٹا:

(i) درست اصولوں کی پیروی کرتا ہے (مثلاً، صرف درست کھیلوں کے نام)

(ii) اقدار کی سطح پر رکھتا ہے (مثلاً، وہ نمبر جہاں نمبر متوقع ہوں)

(iii) جعلی یا احقانہ نہیں ہے

24- ڈیٹا کلیٹنگ کرنے کے طریقوں کی وضاحت کریں۔

جواب: ایک بار جب ہمیں غلطیاں مل جائیں، تو ہمیں انہیں ٹھیک کرنے یا ہٹانے کی ضرورت ہے۔ اسے ڈیٹا کلیٹنگ کہتے ہیں۔ ڈیٹا صاف کرنے

کے لیے درج ذیل چند بنیادی اقدامات ہیں:

- سپیلنگ کی غلطیاں درست کریں: "Football" → "Footbell"
- فارمیٹس کو معیاری بنائیں: "Cricket" → "CRICKET"
- ڈیپلیکیشن ہٹائیں: ایک ہی طالب علم کے لیے بار بار اندراجات کو حذف کریں۔
- گم شدہ اقدار کو ہینڈل کریں: شخص سے دوبارہ پوچھیں یا قطار ہٹا دیں۔
- غلط ڈیٹا ٹھیک کریں: اگر فریکوئنسی دس سو لکھی ہے تو اسے ٹھیک کریں یا ہٹا دیں۔

25- ڈیٹا کے تجزیہ کے لیے آسان ٹولز کے نام بتائیں۔

جواب: آپ Microsoft Excel، Google Sheets یا نوٹ بک میں دستی تحریر جیسے آسان ٹولز سے ڈیٹا کا تجزیہ شروع کر سکتے ہیں۔

26- ڈیٹا کی تصدیق کی وضاحت کریں۔

جواب: لوگ عام طور پر نمبروں سے بہتر آواز پر کوکھتے ہیں، یہی وجہ ہے کہ ہم ڈیٹا ویڈیو لائزیشن کا استعمال کرتے ہیں۔

یاد رکھیں کہ پچھلے مراحل میں، ہم نے پایا کہ کرکٹ سے متعلق کھیلی تھا اس مرحلے میں، ہم ایک چارٹ بنائیں گے جو اسے واضح طور پر

دکھاتا ہے۔ اس سے ہماری دریا فتوں کو ہم جماعتوں، اساتذہ یا دوسروں کے سامنے پیش کرنا آسان ہو جاتا ہے۔

27- ڈیٹا ویڈیو لائزیشن کی کیا اہمیت ہے؟

جواب: ڈیٹا ویڈیو لائزیشن اہم ہے کیونکہ یہ

(i) ڈیٹا کو کبھی آسانی پیدا کرتا ہے،

(ii) ایک نظر میں پیٹرن دکھاتا ہے،

(iii) پیشکشوں یا رپورٹوں میں خیالات کو واضح طور پر پیش کرنے میں مدد کرتا ہے۔

ویڈیو لائبریشن کے لیے ٹولز

5.3

28- Google Sheet کا استعمال کرتے ہوئے چارٹ بنانے کے اقدامات لکھیں۔

جواب: Google Sheets میں چارٹ بنانے کے اقدامات:

- اپنا ڈیٹا قطاروں اور کالموں میں درج کریں۔
- ڈیٹا رینج کو منتخب کریں۔
- 'Insert' مینو پر کلک کریں → 'Chart' کو منتخب کریں۔
- چارٹ کی قسم منتخب کریں (بار، پائی، لائن)۔
- چارٹ کو عنوان، لیبل اور رنگ کے ساتھ حسب ضرورت بنائیں۔

29- اسپریڈ شیٹ سافٹ ویئر کے کیا فوائد ہیں؟

جواب: ہم اسپریڈ شیٹس کا استعمال اس لیے کرتے ہیں کیونکہ وہ:

(i) استعمال میں آسان ہیں

(ii) انٹرایکشن کی ضرورت نہیں (Google Sheets کے لیے)

(iii) چھوٹے سے درمیانے درجے کے منصوبوں کے لیے اچھے ہیں۔

30- حقیقی زندگی میں DSLC کے استعمال کی وضاحت کریں۔

جواب: معیاری 6 DSLC اہم مراحل پر مشتمل ہے، تاہم، حقیقی دنیا کی بعض اسپریڈ شیٹس / منظر ناموں میں، کچھ اضافی مراحل شامل کیے جاسکتے ہیں۔

ایسا ہی ایک مرحلہ ڈیٹا کو ذخیرہ کرنا اور منظم کرنا ہو سکتا ہے، خاص طور پر بڑے منصوبوں یا ٹیم سیکٹرز میں جہاں ڈیٹا میں اور منظم ذخیرہ اہم ہوتے

ہیں۔ یہ عام طور پر DBMS اور SQL جیسے سسٹمز کے ذریعے منظم کیے جاتے ہیں۔

ڈیٹا بیس اور ڈیٹا کو محفوظ کرنا

5.4

31- ڈیٹا بیس کی تعریف کریں۔

جواب: ڈیٹا بیس ڈیٹا کا ایک منظم ذخیرہ ہے۔ یہ ڈیٹا کو ٹیبلوں کی شکل میں ذخیرہ کرتا ہے۔ ہر ٹیبل قطاروں اور کالموں والی ایک چھوٹی اسپریڈ شیٹ کی طرح

ہوتی ہے۔ ڈیٹا بیس ڈیٹا کو مؤثر طریقے سے ذخیرہ کرنے، بازیافت کرنے اور اپ ڈیٹ کرنے میں مدد کرتے ہیں۔

32- ڈیٹا بیس کا مقصد کیا ہے؟

جواب: ڈیٹا بیس کا مقصد:

ایک گندی سرور سے ٹیبل کو صاف کرنے کے بعد، صاف شدہ ڈیٹا کو مناسب شکل میں ذخیرہ کرنے کی ضرورت ہے (یا تو اسپریڈ شیٹ یا ڈیٹا بیس

میں)، تاکہ یہ منظم، قابل تلاش اور تجزیہ کرنے کے لیے تیار ہو۔ منظم ڈیٹا کو ذخیرہ کرنے کے دو اہم طریقے ہیں:

(i) اسپریڈ شیٹس

(ii) ڈیٹا بیس

33- DBMS کی تعریف کریں۔ مثال دیں۔

جواب: ڈیٹا بیس منجمنٹ سسٹم (DBMS) ایک سافٹ ویئر ہے جو ڈیٹا بیس بنانے، ان کا نظم کرنے اور ان کے ساتھ تعامل کرنے کے لیے استعمال ہوتا ہے۔

مثالیں: MySQL، SQLite، اور Microsoft Access۔

34 SQL میں استعمال ہونے والے کچھ استفسارات کی وضاحت کریں۔

جواب: SQL کے کچھ بنیادی استفسارات درج ذیل ہیں:

SELECT * FROM Sports

یہ سادہ تیوری (Query) Sports نامی ٹیبل سے تمام ڈیٹا دکھاتی ہے۔

SELECT Name FROM Sports WHERE Favorite_Sport = 'Cricket'

یہ تیوری Sports ٹیبل میں موجود ان طلباء کے نام دکھاتی ہے جو کرکٹ پسند کرتے ہیں۔

INSERT INTO Sports (Name, Favorite_Sport) VALUES ('Ali', 'Football')

یہ تیوری ایک نیا اندراج (entry) شامل کرتی ہے۔

یہ بنیادی SQL انٹیکس بنے ڈیٹا بیس سے مخصوص ڈیٹا کو مؤثر طریقے سے منظم کرنے اور بازیافت کرنے میں مدد کرتی ہیں۔

حل مشقی سوالات

کثیر الانتخابی سوالات

- i- ڈیٹا سائنس لائف سائیکل (DSL) میں پہلا مرحلہ کیا ہے؟
(الف) ڈیٹا اکٹھا کرنا (ب) ڈیٹا کی بصری پیشکش کرنا (ج) مسئلہ کو سمجھنا (د) ڈیٹا کا تجزیہ کرنا
- ii- بہت سے طالب علموں کی آراء اکٹھا کرنے کا بہترین طریقہ کون سا ہے؟
(الف) IoT آلات (ب) سروے (ج) ویب سکرپٹنگ (د) ڈیٹا مائننگ
- iii- ڈیٹا میں جانبداری (Bias) کیا ہے؟
(الف) ڈیٹا کا بہت بڑا سائز (ب) غیر منصفانہ یا ایک طرف ڈیٹا (ج) اچھی طرح اکٹھا شدہ ڈیٹا (د) تیز رفتار ڈیٹا
- iv- چارٹ بنانے کے لیے ابتدائی افراد کے لیے مثالی ٹول کون سا ہے؟
(الف) Tableau (ب) Python (ج) Google Sheets (د) Hadoop
- v- ایک ڈیٹا میں معلومات کو _____ کی شکل میں محفوظ کرتا ہے۔
(الف) سلائڈز (ب) ٹیبلو (ج) بلاکس (د) لیبرز
- vi- درج ذیل چارٹس میں سے کون سا گروہوں کا موازنہ کرنے کے لیے بہترین ہے؟
(الف) پانی چارٹ (ب) بار چارٹ (ج) لائن چارٹ (د) ایریا چارٹ
- vii- ڈیٹا کا تجزیہ کرنے کا بنیادی مقصد کیا ہے؟
(الف) پرانا ڈیٹا حذف کرنا (ب) ڈیٹا کو سمجھنا (ج) پیٹرنز تلاش کرنا اور فیصلے کرنا (د) انٹرنیٹ پر اپ لوڈ کرنا
- viii- درج ذیل میں سے کون سا ٹول ایک جدید ڈیٹا ویژولائزیشن پلیٹ فارم ہے؟
(الف) MS Paint (ب) Google data Studio (ج) Note Pad (د) WhatsApp
- ix- ڈیٹا کو ذخیرہ کرنے سے پہلے کیا کرنا چاہیے؟
(الف) انٹیمیشنز شامل کرنا (ب) اے کیمن اور تصدیق کرنا (ج) اے کیمن کرنا (د) اس کا ترجمہ کرنا
- x- درج ذیل میں سے کون سی ایک ڈیٹا سوریس کی مثال ہے؟
(الف) یوٹیوب (ب) سروے فارم (ج) وائٹ بورڈ (د) سلائڈشو

مختصر سوالات

- 1- ڈیٹا اکٹھا کرنے سے پہلے مسئلہ کو سمجھنے کی کیا اہمیت ہے؟
جواب: اس سے پہلے کہ ہم ڈیٹا اکٹھا کرنے کے بارے میں بات شروع کریں، مزید غور مسئلے کو واضح طور پر سمجھنا ضروری ہے۔ اگر ہم مسئلے کو صحیح طریقے سے نہیں سمجھتے ہیں، تو ہم غلط ڈیٹا اکٹھا کر سکتے ہیں جس کے نتیجے میں بیکار یا غلط نتائج برآمد ہوں گے۔
- 2- ڈیٹا اکٹھا کرنے کے کوئی دو طریقوں کے نام بتائیں اور ہر ایک کی ایک مثال دیں۔
جواب: سروے: ہم Google فارم کا استعمال کرتے ہوئے سوالات کے ساتھ ایک سروے بنا سکتے ہیں اور جوابات کے لیے اس کا لنک طلبہ کے ساتھ شیئر کر سکتے ہیں۔
APIs: ایک اسکول روزانہ حاضری کے ڈیٹا کو ایک رپورٹ میں پل کرنے کے لیے حاضری API استعمال کر سکتا ہے۔
- 3- "ڈیٹا کلیٹنگ" سے ہمارا کیا مطلب ہے؟
جواب: یاد رکھیں کہ پہلے ہم نے سروے اور دیگر طریقوں کا استعمال کرتے ہوئے ڈیٹا اکٹھا کیا تھا۔ اب، ہمیں یہ جانچنے کی ضرورت ہے کہ آیا وہ ڈیٹا درست اور مکمل ہے۔ اس عمل کو ڈیٹا کلیٹنگ کہتے ہیں۔ یہ قدم اس بات کو یقینی بناتا ہے کہ ہم بعد میں جس ڈیٹا کا تجزیہ کریں وہ قابل اعتبار ہو۔
- 4- ڈیٹا استعمال کرنے سے پہلے اسے تصدیق (validate) کرنا کیوں ضروری ہے؟
جواب: ڈیٹا توثیق اس لیے اہم ہے کیونکہ یہ جانچتی ہے کہ آیا ڈیٹا:
(i) درست اصولوں کی پیروی کرتا ہے (مثلاً، صرف درست کھیلوں کے نام)
(ii) اقدار کی صحیح قسم رکھتا ہے (مثلاً، وہ نمبر جہاں نمبر متوقع ہوں)
(iii) جعلی یا احتمالاً نہیں ہے۔
- 5- چارٹس کا استعمال کرتے ہوئے ڈیٹا کو بصری شکل دینے کا کیا مقصد ہے؟
جواب: لوگ عام طور پر نمبروں سے بہتر تصاویر کو سمجھتے ہیں، یہی وجہ ہے کہ ہم ڈیٹا ویژولائزیشن کا استعمال کرتے ہیں۔ ڈیٹا ویژولائزیشن اہم ہے کیونکہ یہ:
(i) ڈیٹا کو سمجھنے میں آسانی پیدا کرتا ہے،
(ii) ایک نظر میں پیٹرن دکھاتا ہے،
(iii) پیشکشوں یا رپورٹوں میں خیالات کو واضح طور پر شیئر کرنے میں مدد کرتا ہے۔
- 6- ڈیٹا میں جانبداری نتائج کو کیسے متاثر کرتی ہے؟
جواب: جانبداری کا مطلب ہے کہ ڈیٹا سچائی کا صرف ایک حصہ دکھاتا ہے اور مکمل تصویر سے محروم رہ جاتا ہے۔
مثال: اگر ہم صرف لڑکوں سے ان کے پسندیدہ کھیلوں کے بارے میں سروے کرتے ہیں، تو ہم لڑکیوں کی رائے سے محروم رہ جائیں گے۔ نتائج غیر منصفانہ یا لڑکوں کی طرف متعصب ہوں گے۔
- 7- ایک بے ترتیب سروے کو کلین کرنے کے لیے کوئی سے دو مراحل/اقدامات لکھیں۔
جواب: ڈیٹا صاف کرنے کے لیے درج ذیل چند اقدامات ہیں:
• فارمیٹس کو معیاری بنائیں: "Cricket" → "CRICKET"
• ڈیپلیکیشن ہٹائیں: ایک ہی طالب علم کے لیے بار بار اندراجات کو حذف کریں۔

8- چارٹ بنانے کے لیے استعمال ہونے والے کوئی سے دو ٹولز کے نام بتائیں۔

جواب: Microsoft Excel اور Google Sheets ٹیبلز سے چارٹ اور گراف بنانے کے لیے آسان لیکن طاقتور ٹولز ہیں۔

تفصیلی سوالات

- 1- ڈیٹا سائنس لائف سائیکل (DSL) کے چھ مراحل کو ہر ایک کے لیے ایک سطر میں بیان کریں۔
جواب: ڈیٹا سائنس ایک ہی مرحلے میں مکمل نہیں ہوتی۔ یہ اقدامات کا ایک سلسلہ ہے جو کسی مسئلے کو حل کرنے میں مددگار ہوتے ہیں۔ اسے ہم ڈیٹا سائنس لائف سائیکل (DSL) کہتے ہیں۔ DSL عام طور پر مندرجہ ذیل مراحل پر مشتمل ہوتا ہے:
مسئلے کو سمجھنا: زیر غور مسئلے کو واضح طور پر سمجھنا ضروری ہے۔ اگر ہم مسئلے کو صحیح طریقے سے نہیں سمجھتے ہیں، تو ہم غلط ڈیٹا اکٹھا کر سکتے ہیں۔
ڈیٹا اکٹھا کرنا: ڈیٹا خام مال کی طرح ہے، اور ہمیں اسے صاف کرنے، ذخیرہ کرنے یا تجزیہ کرنے سے پہلے صحیح قسم کا ڈیٹا اکٹھا کرنا ہوگا۔ مسئلے اور دستیاب ٹولز کے لحاظ سے ڈیٹا اکٹھا کرنے کے بہت سے طریقے ہیں۔
ڈیٹا کی صفائی اور توثیق: ہمیں یہ جانچنے کی ضرورت ہے کہ آیا وہ ڈیٹا درست، مکمل اور استعمال کے لیے تیار ہے۔ یہ قدم اس بات کو یقینی بناتا ہے کہ ہم بعد میں جس ڈیٹا کا تجزیہ کریں وہ قابل اعتبار ہو۔
ڈیٹا کا تجزیہ کرنا: ڈیٹا کے تجزیہ کا مطلب ہے ڈیٹا کو احتیاط سے جانچنا تاکہ (i) اس کا خلاصہ کیا جاسکے، (ii) مختلف گروپوں کا موازنہ کیا جاسکے، (iii) پیٹرن تلاش کیے جاسکیں۔
ڈیٹا کی تشریح کرنا: ہر حلقہ ڈیٹا سے معنی تلاش کرتا ہے (یعنی رجحانات اور پیٹرن کی نشاندہی کرنا)۔
نتیجہ کی ویڈیو لائزیشن اور ان کا اطلاق:
اس مرحلے میں، ہم ایک چارٹ بنائیں گے جو اسے واضح طور پر دکھاتا ہے۔ اس سے ہماری دریا فتوں کو ہم جماعتوں، اساتذہ یا دوسروں کے سامنے پیش کرنا آسان ہو جاتا ہے۔
- 2- بار چارٹ، پائی چارٹ اور لائن چارٹ کا مثالوں کے ساتھ تفصیلی موازنہ کریں۔
جواب:

خصوصیت / پہلو	بار چارٹ	پائی چارٹ	لائن چارٹ
بنیادی مقصد	مختلف زمروں یا گروپس کا موازنہ کرنا۔	کسی کل کے تناسب یا فیصد کو دکھانا (کل کے حصے)۔	مسلل مدت (عام طور پر وقت) کے دوران رجحانات، پیٹرن، یا تبدیلیاں دکھانا۔
ڈیٹا کی قسم	زمرہ جاتی یا بچر ڈیٹا۔	زمرہ جاتی ڈیٹا، جہاں زمرے کل میں جمع ہوتے ہیں۔	مسلل ڈیٹا، خاص طور پر ٹائم سیریز ڈیٹا۔
محور	ایک محور زمرے دکھاتا ہے، دوسرا اقدار دکھاتا ہے۔	کوئی محور نہیں یہ ایک دائرہ ہے جو سلائسز میں تقسیم ہے۔	دونوں محور مسلسل متغیر دکھاتے ہیں (مثلاً X- محور پر وقت، Y- محور پر قدر)۔
بصری نمائندگی	مستطیل سلائس جن کی لمبائی ان کی نمایندگی کردہ اقدار کے متناسب ہوتی ہے۔	سلائسز میں تقسیم ایک دائرہ، جہاں ہر سلائس کا زاویہ یا رقبہ اس کی مقدار کے متناسب ہوتا ہے۔	سیدھی لائنوں سے منسلک ڈیٹا پوائنٹس۔
بہترین استعمال کی صورت	مختلف مصنوعات کی فروخت کے اعداد و شمار، مختلف شہروں کی آبادی وغیرہ کا موازنہ کرنا۔	مختلف شعبوں کے مارجن، بجٹ کی تقسیم، سروے کے فیصد وغیرہ دکھانا۔	ایک سال میں اشیا کی قیمتیں، ایک ہفتے میں درجہ حرارت میں تبدیلیاں وغیرہ دکھانا۔

طاقتیں	انفرادی ذمہ داری کو بڑھانے اور ان کا موازنہ کرنے میں۔ فرق دکھانے کے لیے اچھا۔	ایک نظر میں سادہ تناسب کی شرحیں دکھانے کے لیے بہترین۔
کمزوریاں	اگر بہت زیادہ زمرے ہوں تو بے ترتیب ہو سکتا ہے۔	اگر بہت سی چھوٹی سلسلے ہوں تو موازنہ کرنا مشکل۔ وقت کے ساتھ رجحانات نہیں دکھا سکتا۔
مثال	پانچ مختلف ممالک کی آبادی کا موازنہ: India, USA, Brazil, China اور Nigeria۔	اسارٹ فون آپریٹنگ سسٹمز کا مارکیٹ شیئر دکھانا: Android, iOS, Others۔

3- DBMS اور SQL کے درمیان فرق کریں۔

جواب:

خصوصیت / پہلو	DBMS (ڈیٹا میں منجست سسٹم)	SQL (اسٹرکچرڈ کوئری لینگویج)
بنیادی تعریف	یہ ایک سافٹ ویئر سسٹم ہے جو ڈیٹا کو منظم کرنے کے لیے استعمال ہوتا ہے۔	یہ ایک پروگرامنگ لینگویج ہے جو ڈیٹا بیس کے ساتھ بات چیت کرنے اور ان میں ہیرا پھیری کرنے کے لیے استعمال ہوتی ہے۔
نوعیت	یہ ایک سافٹ ویئر ایپلیکیشن یا ایک نظام ہے (مثلاً، Oracle, MySQL)۔	یہ ایک کوئری لینگویج یا ایک نحو ہے جو کام انجام دینے کے لیے استعمال ہوتی ہے۔
کام	یہ ڈیٹا میں اور صارفین یا ایپلیکیشنز کے درمیان ایک انٹرفیس کا کام کرتا ہے۔	یہ ڈیٹا کو بازیافت کرنے، داخل کرنے یا اپ ڈیٹ کرنے جیسے مخصوص احکامات پر عمل درآمد کرنے کے لیے ایک ٹول کا کام کرتا ہے۔
اجزاء	یوٹیلٹیز، منظمین، اسٹوریج مینجیور، اور SQL کمانڈز چلانے کے لیے ایک کوئری پروسیسر شامل ہے۔	کمانڈز، شقوں اور آپریٹرز پر مشتمل ہے مثلاً، WHERE, INSERT, SELECT۔
مقصد	سیکورٹی اور سالمیت کو یقینی بناتے ہوئے ڈیٹا کو مؤثر طریقے سے ذخیرہ کرنا، بازیافت کرنا اور ان کا نظم کرنا۔	DBMS میں ذخیرہ شدہ ڈیٹا پر مخصوص کارروائیاں انجام دینا۔

4- طالب علم کی میز سے وہ تمام ڈیٹا حاصل کرنے کے لیے SQL میں ایک کوئری لکھیں جس کا رول نمبر 15 ہے۔

جواب:

SELECT
FROM
WHERE

STUDENT

roll_number = 15;

