主题需求

无法实现：按区域，字段数据缺失

需求1：获取所有学生的平均体测成绩和及格率及每个体测项的平均成绩

维度：

指标：平均成绩，及格率，每个项的平均成绩

涉及字段:height,weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, test\_rank, PE\_score

需求2：获取每个学院的平均体测成绩和及格率

维度：学院

指标：平均成绩，及格率，每个项的平均成绩

涉及字段: student\_number,height,weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, test\_rank, PE\_score

(未做)需求3：获取最高成绩前50的成绩

维度：

指标：每个项的成绩

涉及字段:weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, test\_rank, PE\_score

需求4：获取每个年级的平均体测成绩和及格率

维度：年级

指标：平均成绩，及格率，每个项的平均成绩

涉及字段: student\_number(从学号获取年级),weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, test\_rank, PE\_score

需求5：获取男女学生的平均体测成绩和及格率及每个体测项的平均成绩

维度：性别

指标：平均成绩，及格率，每个项的平均成绩

涉及字段：sex,weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, test\_rank, PE\_score

需求6：获取各地区学生的平均体测成绩和及格率及每个体测项的平均成绩

（通过身份证前几位获得省份）

维度：地区

指标：平均成绩，及格率，每个项的平均成绩

涉及字段：id\_card,weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, test\_rank, PE\_score

需求7：获取每个班级的平均体测成绩和及格率及每个体测项的平均成绩

维度：班级

指标：平均成绩，及格率，每个项的平均成绩

涉及字段：class\_name,weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, test\_rank, PE\_score

(没做)需求8：获取每个年龄段的平均体测成绩和及格率及每个体测项的平均成绩

维度：年龄

涉及字段：id\_card,weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, test\_rank, PE\_score

总结：

涉及维度：

学院，班级，地区，性别，年级，年龄

涉及指标：

平均成绩，及格率，每个项的平均成绩，每个项的成绩

涉及表：

事实表：student\_body\_test\_2020

维度表：无

涉及字段：

学院，年级维度：student\_number

班级维度：class\_name

地区维度：id\_card

性别维度：sex

指标字段：height,weight,vital\_capacity,meter\_50,jump, sit\_and\_reach, meter\_800, meter\_1000, sit\_up, pull\_up, left\_vision, right\_vision, height\_weight\_score, vital\_capacity\_score, meter\_50\_score, jump\_score, sit\_and\_push\_score, meter\_800\_1000\_score, sit\_and\_reach\_pull\_up\_score, total\_score, , PE\_score

区分合格与否：test\_rank=”合格”即为合格

需要清洗数据：不需要

需要转换字段：

student\_number转换为department,grade

id\_card转换为地区，年龄