附件1

选题具体要求

本次参赛论文分为“统计建模类”、“大数据应用类”、“市场调查分析类”和“生物医学类”。“统计建模类”大赛旨在揭示批量数据背后的因素，诠释社会经济现象，对经济和社会发展做出预测或判断。“大数据应用类”大赛旨在通过大赛构建良好的资源共享平台，促进大数据与云计算等战略新兴产业的融合，促进大数据专业技术人才的培养。“市场调查分析类”大赛在于分析市场情况，了解市场现状及其发展趋势，为市场预测和营销决策提供客观的、正确的资料。“生物医学类”大赛旨在通过生物标本、动物实验、人类个体和群体的数据分析，回应对公众和社会关注的生命科学和人类健康问题。

“统计建模类”选题可以是与社会发展息息相关，或与政府统计部门及统计热点工作相关的经济、贸易、金融、生态、人口、交通、能源、城市化、农业、生物、医药、卫生等诸多方面的问题，本次大赛特别建议以“高质量发展”、“大数据”和“数字经济”作为关键词考虑选题。由参赛者自行搜集数据，提出问题和假设条件，建立模型，运用统计分析方法和统计分析软件进行模型求解，阐明主要结论及现实意义，并对结果进行分析与检验，讨论模型的优缺点和改进方向。

“大数据应用类”选题是根据公司、企业和政府统计部门、科研院所等方面在生产和研究过程中遇到的实际问题，围绕“数据”、“算法”和“应用”，经过适当简化加工提炼出所需解决的问题，要求参赛者能够利用相关的软件从网页中爬取所需的数据，通过统计模型运用进行数据分析（特别是文本数据）的处理，从批量的数据中挖掘有效信息建立统计模型解决实际问题，要求参赛者对相关领域背景问题能有足够的理解、能够做到模型评价的客观性、创新性和准确性。

“市场调查分析类”选题是为了结合社会的热点、企业的需求和个人的关注方向，围绕“市场调查”、“数据分析”和“模型预测”，由参赛者提出需要解决的问题，制定调查方案，通过调查问卷收集所需数据，针对所提出的问题分析数据，给出解决问题的对策和建议。

“生物医学类”选题主要围绕生物科学、生物技术、基础医学、老年医学、预防医学、医学影像等领域的理论和实际需要，提出需要解决的问题，并采用观察性或试验性研究方法获取研究数据，建立模型。要求参赛者有明确的数据来源，对模型的优缺点和局限性的评价要准确、客观，并提出改进意见。