

大数据分析技术在金融衍生品专项审计中的应用

王 强¹, 吴 琼¹, 郭正良²

(1. 中国石油化工集团有限公司境外审计中心, 北京 100029; 2. 中国石化集团国际石油勘探开发有限公司, 北京 100029)

摘要: 国际石油公司通常采用金融衍生品进行套期保值, 规避因价格变动带来的风险, 实现风险对冲。由于金融衍生品的高风险性, 有必要进行审计监督, 使之成为公司风险管理的有效工具。基于 Python 语言, 对金融衍生品数据进行全样本采集、处理和分析, 检查金融衍生品操作方案执行上存在的问题。大数据分析技术能够赋能内部审计高质量发展, 提高了审计的质量和效率, 有助于实现集中统一、权威高效和全面覆盖的审计目标。

关键词: Python; 大数据分析; 金融衍生品; 内部审计

中图分类号: F239.4; F239.45; C37 **文献标志码:** A **文章编号:** 1671-1807(2024)07-0049-08

金融衍生品是一种由基础资产派生出的金融工具, 主要包括期货、远期、互换和期权, 具有定价复杂、高杠杆、高风险、国际化程度高等特征。由于金融衍生品市场具备价格发现和风险规避的功能, 国际石油公司通常利用衍生品进行套期保值, 实现对生产经营风险的管控。根据标的资产性质的不同, 中石化将金融衍生品业务划分为商品类衍生品业务和货币类衍生品业务。商品类套期保值是指为管理实货价格随市场情况变化导致实货面临盈亏变动风险而开展的, 以管理采购成本、销售收入、库存价格等, 实现风险对冲目的, 与实货品种相关, 与风险敞口数量相当, 与面临风险方向相反, 与期货期限相匹配的商品类套期保值业务。套期保值作为一种风险控制手段, 具有高度的风险性, 存在基差风险、信用风险、决策风险、财务风险和流动性风险。因此, 对金融衍生品业务开展情况进行审计监督, 使金融衍生品业务真正成为公司风险管理的工具。对金融衍生品进行内部审计可以达到以下目的: ①内部审计可以帮助企业识别、评估和管理风险, 确保衍生品投资活动的合规性和风险控制; ②金融衍生品的交易涉及大量的资金流动和复杂的交易结构, 内部审计对衍生品交易的内部控制进行审查, 确保交易的透明度和规范性, 防止内部人员利用衍生品进行不当操作; ③金融衍生品的会计处理和信息披露对于企业的财务报告质量至关重

要, 内部审计可以对衍生品的会计处理和信息披露进行审查, 确保其符合会计准则和相关法规, 提高财务报告的质量和准确性; ④金融衍生品的交易受到严格的监管和法规限制, 内部审计可以帮助企业确保其衍生品交易活动符合相关法规和监管要求, 避免因违规操作导致的法律风险和声誉风险; ⑤通过对衍生品交易的风险、成本、收益等进行综合分析, 内部审计可以为管理层提供关于金融衍生品交易的决策支持和建议, 确保企业的战略目标和业务发展。

开展金融衍生品业务专项审计十分必要。审计对套期保值交易发挥正向促进作用, 没有挫伤国有企业正常的套期保值需求^[1-2]。金融衍生品监管虽增加了审计费用, 但提高了国企审计质量, 并降低了国企被出具非标准审计意见的概率^[3]。金融衍生品业务审计侧重于检查交易品种是否与主业经营产品相关, 是否超越规定的经营范围, 年度保值规模和风险指标是否超标, 是否建立科学合理的激励约束机制^[4]。评估套保效果时应将期货端与现货端结合起来分析, 应重点关注金融衍生品交易内控制度设计和运行的有效性^[5]。开展合规审计应对市场风险, 可靠性审计应对信用风险, 全面审计应对流动性风险和操作风险, 风险导向性审计应对法律风险^[6]。大数据分析技术在金融衍生品审计中发挥关键作用。向大数据要资源, 向信息化要效率, 大

收稿日期: 2023-12-26

作者简介: 王强(1989—), 男, 安徽安庆人, 博士, 高级工程师, 研究方向为大数据审计、审计理论与实务; 吴琼(1984—), 女, 河北石家庄人, 硕士, 高级经济师, 研究方向为内部审计、审计理论与实务; 通信作者郭正良(1988—), 男, 河北张家口人, 硕士, 高级工程师, 研究方向为油气田开发工程、渗流力学。

数据处理和分析技术掀起了新一轮审计思维变革的浪潮^[7],传统的审计方法已经难以应付爆炸式增长的海量数据,数字化转型和大数据审计已成为审计转型升级的必然趋势^[8],利用大数据技术来提高审计效率和准确性,更好地为企业可持续高质量发展保驾护航^[9]。深化技术创新审计在各领域实践应用,加强研究大数据时代下技术创新审计信息化建设^[10]。大数据审计是对传统审计内涵的一次重大拓展,也是对审计模式的一次重大变革,让大数据渗透到审计工作的规划、实施、运行和评价的各个环节和层面^[11]。

大数据技术已经在审计领域中得到充分应用,但在金融衍生品审计中的应用相对较少。学者们对金融衍生品审计的必要性、重点查证内容、大数据审计理论和实践均展开了丰富的研究,但对于大数据技术背景下金融衍生品业务审计的理论方法、查证步骤和实施路径仍然缺乏深入研究和有效探索。因此,在对被审计单位开展商品类衍生品业务专项审计时,利用 Python 查证金融衍生品操作方案执行是否规范。以大数据全样本审计为基础,重点审查交易执行和规范操作情况,将操作方案中的交易策略核心内容与实际衍生品交易执行情况进行对比,查看是否存在不匹配情况,所有实际交易的业务逻辑是否都在操作方案的市场分析和交易策略中予以描述等。探究大数据审计在金融衍生品审计中的作用,借助全样本审计揭示全面覆盖风险,以期能为金融衍生品审计与大数据技术相融合的信息化和智能化审计道路提供助力。

1 重点查证内容

金融衍生品业务涉及多个部门,前台的业务部门制订年度计划并执行,开展实货交易管理、期纸货交易管理等;中台的风险控制执行部门主要进行操作方案管理、业务授权管理、交易盯市盈亏,以及交易对手管理等;后台的财务部门关注交易的执行、资金结算、套保效果评价,以及套保会计核算

等。金融衍生品主要业务流程和部门职责如图 1 所示。

根据衍生品前、中、后台业务流程和重要节点,可以从事前管理、事中管理和事后管理 3 个阶段梳理审计重点关注内容。

1.1 事前管理阶段,关注计划制订和执行

审查金融衍生品年度计划目标、套保策略、风险敞口等;审查金融衍生品业务年度计划是否包括年度实货经营和保值规模、套期保值策略、最大净持仓规模、止损限额、保证金预警线等指标。

1.2 事中管理阶段,关注操作方案管理、盯市盈亏、交易对手管理和权限管理

(1)操作方案管理。审查操作方案的完整性和合理性;操作方案是否明确风险敞口、总体和阶段套保目标、套保逻辑等;操作方案的审批是否合规,交易执行和规范操作情况。

(2)盯市盈亏。重点关注持仓限额、保证金预警线和止损限额等风险指标设置是否按规定进行审批,风险指标统计口径和计算逻辑是否合规,能否真实准确地反映套保综合效果。

(3)交易对手管理。重点关注对交易对手是否进行审慎全面的尽职调查和风险评估、违约或破产等突发事件是否立即上报处理等。

(4)授权管理。是否建立授权审批权限机制,前中后台岗位设置是否相容,授权书是否明确交易权限的人员名单、交易品种和额度。

1.3 事后管理阶段,关注报告制度、事后评价机制和套期保值会计

(1)报告制度。重点审查风险日报、保证金日报、月度和季度统计报表,月度对账资料以及年度专项报告等,是否按制度流程及时、准确、完整上报,是否存在错报、漏报或瞒报情况。

(2)事后评价机制。重点审查是否对每个操作方案进行合理评估,是否完整地交易价格、盈亏情况、套保比例和套保效果进行描述。



图 1 金融衍生品主要业务流程和部门职责

(3)套期保值会计。结合企业会计准则,审查企业应用套期保值会计的有效性、合理性和准确性。

2 数据采集与处理

以操作方案管理为重点研究内容,开展操作方案执行情况审查。根据被审计单位金融衍生品管理办法规定,开展金融衍生品业务前,应根据年度计划和套保策略制定具体操作方案;操作方案履行相应的审批程序后,方可开展操作;应确保操作方案逻辑合理、审批及时、执行规范,应对操作方案的执行、调整等情况进行跟踪,定期开展实际执行与操作方案的一致性检查。因此,在金融衍生品专项审计中,需要针对操作方案执行情况进行检查。

在利用 Python 开展数据分析时,需先下载和导入相应的第三方库,如 pandas、re、numpy 等。pandas 是基于 numpy 的一种工具,为解决数据分析任务而创建的。pandas 融入了大量库和一些标准的数据模型,提供了高效地操作大型数据集所需的工具,并提供了大量能够快速便捷地处理数据的函数和方法。正则表达式描述了一种字符串匹配

的模式,用来检查一个字符串是否含有某种子串、将匹配的子串替换或者从某个串中取出符合某个条件的子串等。re 是 Python 独有的匹配字符串的模块,该模块基于正则表达式,对于字符串进行模糊匹配找到想要的内容信息。

所分析的案例中,“hedge plan”表即操作方案明细表,包含了操作方案编号、衍生品产品名称、计划交易起止日期、数量上下限、操作方案说明、审批时间和审批人等信息。“paper”表即衍生品成交明细表,包含了衍生品编号、所属操作方案编号、买卖方向、成交日期、合约月、成交价格、成交数量、交易员和实货信息等。以操作方案编号将两张表进行左连接,即可得到每个操作方案项下的衍生品成交信息,并与操作方案规定要素进行对比,确定是否有未按操作方案规定内容开展衍生品交易的问题。

先导入 pandas、re、datetime、numpy 库,再读取操作方案明细表和衍生品成交明细表,以衍生品成交明细表为左表,将两张表合并。Python 代码如下。

```
import pandas as pd
import re
import datetime
import numpy as np
hedge_plan=pd.read_excel(r./金融衍生品审计.xlsx',sheet_name='hedge plan')
paper=pd.read_excel(r./金融衍生品审计.xlsx',sheet_name='paper')
df=paper.merge(hedge_plan,left_on='Hedge Plan',right_on='套保计划编号')
```

合并后的新表共有 4 213 行数据,每行数据包含 29 列信息,表 1 列示了前 20 行衍生品序号、交易日期、衍生品名称、操作方案编号、成交价格和数量等信息

(编号和数据经过脱密处理)。后续审计查证均是在该表、操作方案明细表和衍生品成交明细表的基础上进行数据的采集、处理、分析和保存。

表 1 金融衍生品操作方案和交易明细表

交易编号	成交日期	纸货交易日期起	纸货交易日期至	品种	操作方案编号	数量/桶
*****/2023010073	2022-12-30	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010072	2022-12-30	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010071	2022-12-29	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010070	2022-12-29	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010069	2022-12-29	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010068	2022-12-29	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010067	2022-12-29	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010066	2022-12-29	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010065	2022-12-28	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010064	2022-12-28	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010063	2022-12-27	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010062	2022-12-27	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010061	2022-12-27	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010060	2022-12-27	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010059	2022-12-27	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010058	2022-12-27	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010057	2022-12-23	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010056	2022-12-23	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000
*****/2023010055	2022-12-19	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	9 000
*****/2023010054	2022-12-19	2023-01-01	2023-01-31	OTC-ULSD10-D	*****/2301/0001	6 000

3 审计查证步骤

根据审计查证重点, 操作方案执行不规范审计查证主要关注以下 6 个方面: 审批流程是否完善, 是否缺失审批人或审批环节; 是否存在先衍生品交易后操作方案审批; 是否在操作方案规定的计划交易日期范围外开展衍生品交易; 衍生品交易方向是否与操作方案规定的方向相冲突; 衍生品成交时的油价是否在操作方案规定的油价范围之外; 衍生品成交数量是否超过操作方案规定上下限。利用 Python 语言编写代码, 就这 6 方面的问题进行详细分析和查证。

```
clue_hedgep_id=[]
for i in range(hedge_plan.shape[0]):
    approver=re.findall(r'[0-9].?([\u4e00-\u9fa5]+)',hedge_plan[审批人][i])
    if len(approver)==1:
        clue_hedgep_id.append(hedge_plan[套保计划编号][i])
```

运行代码, 发现有 27 份操作方案只完成了一级审批, 审批流程不完整, 如图 2 所示。通过查阅资料和现场访谈, 证实这 27 份操作方案本应某部门经理李某审批, 后李某升任其他分公司副总经理, 该部门经理缺失, 故后续操作方案仅由分管领导一级审批。

3.2 先交易后审批, 在操作方案审批前开展衍生品交易

```
clue_paper_id=[]
for i in range(hedge_plan.shape[0]):
    approve_time=".".join(re.findall(r:'[0-9-:\s]+',hedge_plan[审批时间][i][0]).strip('-').strip())
    approve_time=datetime.datetime.strptime(approve_time,"%Y-%m-%d %H:%M:%S")
    if df[Ctr Date][i]<approve_time:
        print(df[Ctr Date][i]-approve_time)
        clue_paper_id.append(df[Contract No.][i])
```

运行代码, 发现有 14 笔衍生品交易在其操作方案未完成审批的情况下已经成交, 提前天数为 1~43 d, 如图 3 所示。通过进一步查证, 这些操作方案是由海外某子公司制订并审批, 当时已在系统中录入操作方案, 由于分管领导休假不在岗未做好业务授权且线上审批系统无法登录, 直至其返回工作岗位后才在线上完成审批, 交易员在电话通知分管领导并得到口头同意的情况下开展上述交易。

```
*****/2201/0004, *****/2202/0006, *****/2202/0009,
*****/2203/0012, *****/2204/0015, *****/2204/0016,
*****/2205/0018, *****/2206/0021, *****/2207/0029,
*****/2208/0030, *****/2209/0032, *****/2210/0034,
*****/2211/0035, *****/2203/0017, *****/2208/0021,
*****/2208/0022, *****/2208/0023, *****/2210/0029,
*****/2210/0030, *****/2112/0055, *****/2207/0019,
*****/2208/0021, *****/2208/0022, *****/2209/0023,
*****/2209/0024, *****/2210/0029, *****/2210/0030.
```

图 2 审批流程不完整的操作方案编号

3.1 审批流程不完善, 缺失审批人或环节

根据被审计单位金融衍生品管理办法规定, 金融衍生品操作方案需要部门经理、分管领导逐级审批。如果只有一级审批, 表明操作方案审批流程不完善, 缺失审批人或环节。操作方案明细表中存在字段: 审批人或审批时间, 列示了每个操作方案的审批时间和审批人。正则表达式是用于匹配字符串中字符组合的模式, 使用 Python 的正则表达式 re 模块, 提取出每个操作方案的审批人和审批时间, 根据审批人数得到审批节点, 当只有一级审批的操作方案。Python 代码如下。

对比操作方案审批时间和期纸货成交时间, 如果期纸货成交日期要早于操作方案审批时间, 表明该笔交易属于先交易后审批。操作方案明细表中审批时间, 将该字段分解并提取审批时间, 将此时间与纸货交易明细表中的成交日期(ctr date)字段进行比较, 当成交日期早于操作方案审批日期时, 表明该笔交易在操作方案审批前即成交。Python 代码如下。

3.3 未在操作方案规定的计划交易日期范围内开展衍生品交易

操作方案规定了开展期纸货交易的时间, 在实际开展业务时应该严格按照操作方案规定的计划交易日期。因此, 需要审查是否检查未在操作方案规定的计划交易日期范围内开展期纸货交易的问题。操作方案明细表中有以下两个字段: 期纸货交易起始日期和期纸货交易截止日期; 期纸货交易明细表中有期纸

1. 先交易后审批的纸货编号为: ****2022120028, 提前交易的天数为: -12 days +05:16:59
2. 先交易后审批的纸货编号为: ****2022120009, 提前交易的天数为: -28 days +09:07:19
3. 先交易后审批的纸货编号为: ****2022120008, 提前交易的天数为: -43 days +06:15:14
4. 先交易后审批的纸货编号为: ****2023010011, 提前交易的天数为: -5 days +08:42:36
5. 先交易后审批的纸货编号为: ****2023010010, 提前交易的天数为: -16 days +14:34:11
6. 先交易后审批的纸货编号为: ****2022120005, 提前交易的天数为: -11 days +12:19:33
7. 先交易后审批的纸货编号为: ****2023010001, 提前交易的天数为: -18 days +12:01:52
8. 先交易后审批的纸货编号为: ****2023010002, 提前交易的天数为: -1 days +12:11:45
9. 先交易后审批的纸货编号为: ****2023020002, 提前交易的天数为: -3 days +08:33:26
10. 先交易后审批的纸货编号为: ****2023010003, 提前交易的天数为: -15 days +11:52:38
11. 先交易后审批的纸货编号为: ****2023010004, 提前交易的天数为: -15 days +11:52:38
12. 先交易后审批的纸货编号为: ****2022120019, 提前交易的天数为: -15 days +11:52:38
13. 先交易后审批的纸货编号为: ****2023010001, 提前交易的天数为: -3 days +08:04:26
14. 先交易后审批的纸货编号为: ****2022120010, 提前交易的天数为: -29 days +13:40:09

图 3 先交易后审批的衍生品交易

货成交日期(Ctr Date)。通过比较这 3 个字段,检查 Ctr Date 是否处于纸货交易起止日期内。如若,则表

明存在未在操作方案规定的计划交易日期范围内开展纸货交易的问题。Python 代码如下。

```
df_outperiod=df[~((df['Ctr Date']<=df['纸货交易日期起']) && (df['Ctr Date']>=df['纸货交易日期至']))]
df_outperiod['滞后天数']=df_outperiod['Ctr Date']-df['纸货交易日期至']
df_outperiod[['Contract No.', 'Hedge Plan', 'Ctr Date', '纸货交易日期起', '纸货交易日期至', '滞后天数']].sort_values(by=['滞后天数', 'Ctr Date'], ascending=[False, True])
```

共有 14 笔衍生品在操作方案规定的截止日期后成交,滞后天数 1~12 d,见表 2。经过访谈和沟通,原操作方案在预定时间范围内尚未完成操作,后续相关

交易仍沿用前期计划,且后续交易的套保背景及目的与原计划一致,因此未对操作方案进行调整,导致衍生品成交日期晚于操作方案规定的截止日期。

表 2 在操作方案规定的计划交易日期范围之外开展的衍生品交易

序号	交易编号	操作方案编号	成交日期	纸货交易日期起	纸货交易日期至	滞后天数
927	*****2022100054	*****/2210/0171	2022-10-12	2022-08-31	2022-09-30	12
928	*****2022100053	*****/2210/0171	2022-10-12	2022-08-31	2022-09-30	12
929	*****2022100052	*****/2210/0171	2022-10-12	2022-08-31	2022-09-30	12
930	*****2022100051	*****/2210/0171	2022-10-12	2022-08-31	2022-09-30	12
236	*****2022120010	*****/2212/0023	2022-12-09	2022-11-01	2022-11-30	9
237	*****2022120009	*****/2212/0023	2022-12-09	2022-11-01	2022-11-30	9
931	*****2022100050	*****/2210/0171	2022-10-04	2022-08-31	2022-09-30	4
932	*****2022100049	*****/2210/0171	2022-10-04	2022-08-31	2022-09-30	4
2997	*****2022030012	*****/2203/0016	2022-03-01	2022-02-21	2022-02-28	1
3000	*****2022030077	*****/2203/0014	2022-03-01	2022-02-03	2022-02-28	1
3029	*****2022030056	*****/2202/0005	2022-03-01	2022-01-06	2022-02-28	1
3030	*****2022030055	*****/2202/0005	2022-03-01	2022-01-06	2022-02-28	1
3031	*****2022030054	*****/2202/0005	2022-03-01	2022-01-06	2022-02-28	1
3032	*****2022030053	*****/2202/0005	2022-03-01	2022-01-06	2022-02-28	1

3.4 未按照操作方案规定的买卖方向开展交易

期货纸货交易明细表中有期纸货成交时的买卖方向,操作方案明细表中的字段“套保计划说明”有

买入或卖出说明,对比检查纸货是否按照操作方案规定的方向成交。Python 代码如下。

```
clue_paper_id=[]
for i in range(df.shape[0]):
    if df['B/S'][i]=='Buy':
        if '卖' in df['套保计划说明'][i] and '买' not in df['套保计划说明'][i] and '月差' not in df['套保计划说明'][i] and '平回' not in df['套保计划说明'][i]:
            clue_paper_id.append(df['Contract No.'][i])
for i in range(df.shape[0]):
    if df['B/S'][i]=='Sell':
        if '买' in df['套保计划说明'][i] and '月差' not in df['套保计划说明'][i] and '卖' not in df['套保计划说明'][i] and '平回' not in df['套保计划说明'][i]:
            clue_paper_id.append(df['Contract No.'][i])
```

共有 18 笔衍生品成交时违反了操作方案的套保策略,未按规定的买卖方向成交,其中有 6 笔衍生品交易方向为买入,而操作方案规定的方向为卖出,有 12 笔衍生品交易方向为卖出,而操作方案规定的方向为买入,如图 4 所示。例如,编号为“*****/2022110021”的衍生品于 2022 年 10 月 18 日买入,而对应操作方案的计划说明为“...计划卖出最多 8000 吨 11 月纸货,过程中可根据实货当日销售情况进行卖出保值操作。”

3.5 纸货成交时,油价未在操作方案规定的布油价格区间

```
contract_id = []
for i in range(df.shape[0]):
    description=df['Hedge Note Description'][i]
    brent=re.findall(r'(\d+\.\d+)',description)
    brent=map(lambda x:float(x),brent)
    brent=list(filter(lambda x:x>50 and x<150,brent))
    price_range=''.join(re.findall(r'布油首行(\d+-\d+)',df['套保计划说明'][i])).split('-')
    price_range=list(map(lambda x:x if x!='' else int(x),price_range))
    if price_range! =[] and brent:
        price_range_min,price_range_max=price_range
        for j in brent:
            if (j<price_range_min-2 and df['B/S'][i]=='Sell') or (j>price_range_max+2 and df['B/S'][i]=='Buy'):
                contract_id.append(df['Contract No.'][i])
contract_id=list(set(contract_id))
```

共有 67 笔衍生品成交时,布伦特原油价格未在操作方案规定的布油价格区间,其中在价格区间上限买入的纸货有 9 笔,在价格区间下限卖出的纸货有 54 笔,见表 3。2022 年油价波幅巨大,暴涨暴跌

操作方案的套保计划说明中会描述在某个布伦特原油价格区间内开展期纸货交易,如某操作方案的套保计划说明:“...计划在布油首行 65~80 美元区间里,卖出 35~40 万桶...”,期纸货台账中的字段“Hedge Note Description”会记录期纸货成交时的布伦特油价。例如,“Jan IPE 932 Brent 83.97 (Mar 23)”。需要使用正则表达式将这两个字段中的布油价格区间和价格提取出来,并进行对比,确认成交时的布油价格是否在操作方案规定的布油价格区间内。Python 代码如下。

频繁,业务部门部分操作方案中的目标价位未能也难以根据市场变化及时更新,纸货交易操作未严格按照操作方案规定内容开展。

```
*****/2022110021, *****/2022110020, *****/2022040023,
*****/2022040024, *****/2022040025, *****/2022030073,
*****/2022020043, *****/2023080001, *****/2022090010,
*****/2022090009, *****/2022090008, *****/2022090007,
*****/2022090006, *****/2022090005, *****/2022120006,
*****/2022120005, *****/2022120004, *****/2022030140.
```

图 4 与操作方案规定的交易方向不一致的衍生品交易

表 3 未在操作方案规定区间内的衍生品交易

交易编号	套保计划说明	套保计划描述
*****2023010028	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 848.50 Brent (Feb23)77.60
*****2023010027	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 848.50 Brent (Feb23)77.60
*****2023010026	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 822.75 Brent (Feb23)76.28
*****2023010025	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 822.75 Brent (Feb23)76.28
*****2023010024	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 814.5 Brent (Feb23)75.48
*****2023010023	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 814.5 Brent(Feb23)75.48
*****2023010022	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 822 Brent (Feb23)76.68
*****2023010021	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 822 Brent (Feb23)76.68
*****2023010020	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 824.25 Brent (Feb23)77.3
*****2023010019	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 250000 桶	Jan IPE 824.25 Brent (Feb23)77.3
*****2023010015	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 10000 吨	Brent 2月 76.52
*****2023010014	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 10000 吨	Brent 2月 76.25
*****2023010006	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 10000 吨	Brent 2月 76.47
*****2023010005	计划在布油首行 80~100 美元区间里;在区间上沿择机卖出最多 10000 吨	Brent 2月 76.4

3.6 纸货成交数量超过操作方案规定的成交数量范围

统计操作方案项下的纸货成交数量,检查衍生

```
df[数量(带方向)]=0
for i in range(df.shape[0]):
    if df[B/S][i]='Buy':
        df[数量(带方向)][i]=df[Qty][i]
    else:
        df[数量(带方向)][i]=-df[Qty][i]
paper_amount=pd.pivot_table(df,index=[u套保计划编号],values=u数量(带方向),aggfunc=np.sum).reset_index()
paper_amount[数量(带方向)]=paper_amount[数量(带方向)].apply(lambda x:abs(x))
df2=hedge_plan.merge(paper_amount,left_on=套保计划编号,right_on=套保计划编号)
df2[限制及生效条件].fillna(0,inplace=True)
flag,clue_hedgep_id=0,[]
for i in range(df2.shape[0]):
    if not (df2[数量(带方向)][i]>=df2[限制及生效条件][i] and df2[数量(带方向)][i]<=df2[纸货数量上限][i]):
        if df2[数量(带方向)][i]!=0:
            flag+=1
            print(df2[套保计划编号][i],df2[限制及生效条件][i],df2[纸货数量上限][i],df2[数量(带方向)][i])
            clue_hedgep_id.append(df2[套保计划编号][i])
```

共有 6 份操作方案项下的衍生品交易成交数量超过了规定的数量上下限,其中有 4 份低于操作方案规定的数量下限,有 2 份高于操作方案规定的数量上限,如图 5 所示。经过审计组现场确认,前 4 份操作方案在执行时间段内市场没有达到预期价格,交易员一直未等到合适的时机建仓,故在计划计划期内操作方案的数量未足额完成;后 2 份则是由于交易员未认真审核操作方案规定的衍生品成交数量上限,在建仓时发现市场预期良好,根据主观意愿超规模下单导致。

4 结论

通过 Python 语言对审计数据进行全样本的钻取和分析,实现了抽样审计向全样本审计转变,扩大了审计覆盖范围,提升了审计的质量和效率。通过上述对金融衍生品操作方案执行情况审计实例,对应用大数据分析技术开展金融衍生品专项审计,得到以下结论。

(1)深化大数据分析技术在审计中的应用。大

品成交数量是否在操作方案规定的成交数范围之内。Python 代码如下。

数据审计是提升审计覆盖面的重要途径,是拓展审计广度和深度的重要手段。积极向大数据要资源,向信息化要效率,多措并举探索大数据审计新思路、新技术和新方法,充分发挥数据潜能。

(2)积极开展不同层面的沟通协作,努力发掘审计需求,彰显大数据在助力审计项目提速增效上的作用。审计项目组成员需要积极有效的沟通,确定审计的大数据分析需求,由组内大数据审计尖兵确认审计目标的实现路径和方法;不同审计组之间也应该沟通和交流,分享大数据审计值得推荐的查证方法和技术;不同行业的内部审计也应及时沟通和交流,打破技术壁垒,共同促进大数据技术在内部审计中的繁荣发展。

(3)梳理审计查证步骤,探索建立平台式无感化审计。以数据为核心,构建数智化审计,深化远程在线审计应用力度,全面推动抽样审计向全样本审计转变,推动事后监督向实时监督转变,加快业审融合平台式审计转型,实现大数据无感式审计。

操作方案编号为:**/2204/0008 纸货成交数量下限为: 100000,纸货成交数量上限为: 500000,纸货实际成交量为: 70000
 操作方案编号为:**/2207/0022 纸货成交数量下限为: 100000,纸货成交数量上限为: 500000,纸货实际成交量为: 70000
 操作方案编号为:**/2209/0031 纸货成交数量下限为: 50000,纸货成交数量上限为: 300000,纸货实际成交量为: 40000
 操作方案编号为:**/2210/0033 纸货成交数量下限为: 10000,纸货成交数量上限为: 300000,纸货实际成交量为: 5000
 操作方案编号为:**/2201/0001 纸货成交数量下限为: 100000,纸货成交数量上限为: 400000,纸货实际成交量为: 1150000
 操作方案编号为:**/2206/0015 纸货成交数量下限为: 0,纸货成交数量上限为: 20000,纸货实际成交量为: 135000

图 5 纸货成交数量超过上下限的操作方案

参考文献

[1] 刘芳,王美英. 国家审计与国有企业金融衍生品投资研究[J]. 当代财经, 2021(2): 137-148.

[2] 王羽. 运用 GONE 理论提升金融衍生品审计质量研究[J]. 中国内部审计, 2023(12): 80-84.

[3] 赵峰,张甜甜,马光明,等. 金融衍生品监管与企业审计:基于中国金融衍生品监管指数的实证分析[J]. 财贸经济, 2021, 42(3): 77-95.

- [4] 车中扬. 金融衍生品业务审计要点[J]. 中国石化, 2020(8): 54-56.
- [5] 鄢琛. 国有企业金融衍生品审计案例研究[D]. 南京:南京审计大学, 2022.
- [6] 王少华, 袁显朋. 金融衍生品业务风险及审计应对策略[J]. 行政事业资产与财务, 2021(10): 108-110.
- [7] 彭梅, 齐玲, 杨丽, 等. 基于多维信息量化在建设项目招标投标中的审计判别模型[J]. 科技和产业, 2020, 20(3): 136-141.
- [8] 刘仑, 刘娜娜. 企业内部审计数字化转型研究[J]. 现代审计与会计, 2023(12): 24-26.
- [9] 王强. 商业智能技术在融资性贸易审计中的应用[J]. 中国内部审计, 2021(7): 22-27.
- [10] 时迎健, 林向义, 罗杰, 等. 基于知识图谱分析的技术创新审计研究进展与展望[J]. 科技和产业, 2021, 21(8): 209-215.
- [11] 吴小满, 彭聪, 彭乔依, 等. 大数据时代经济责任审计发展趋势及路径探究[J]. 会计之友, 2023(24): 28-34.

The Application of Big Data Analysis Technology in Internal Audit of Financial Derivatives

WANG Qiang¹, WU Qiong¹, GUO Zhengliang²

(1. Overseas Audit Center of SINOPEC, Beijing 100029, China;

2. SINOPEC International Petroleum Exploration and Production Corporation, Beijing 100029, China)

Abstract: International oil companies usually use financial derivatives for hedging to avoid risks caused by price changes and achieve risk hedging. Due to the high risk of financial derivatives, it is necessary to carry on audit to make them an effective tool for risk management. Based on Python language, collection, processing, and analysis of financial derivative data are conducted to check for issues in the execution of financial derivative operation plans. Big data analysis technology can empower the high-quality development of internal auditing, improve the quality and efficiency of auditing, and it is helpful to achieve the audit objectives of centralized, unified, authoritative, efficient, and comprehensive coverage.

Keywords: Python; big data analysis; financial derivatives; internal audit