The sample consists of 1,000 observations from Lebanon, collected as part of an externally funded grant under a broader research initiative focused on entrepreneurship in the MENA region. Data collection employed the Computer-Assisted Personal Interviewing (CAPI) methodology, with the survey instrument structured as a comprehensive questionnaire. The questionnaire covered a wide range of topics, including respondents' demographic characteristics, employment history, educational background, familial ties, social status, household details, financial situation, personality traits, entrepreneurial orientation, entrepreneurial intention, and the broader entrepreneurial environment.

The survey was professionally translated into Arabic to ensure linguistic accuracy and cultural relevance, capturing the nuances of the region. The translated version was programmed using proprietary CAPI platforms that incorporated scripting, data collection, field management, and quality control features, enabling real-time monitoring and data validation. Data collection was conducted using smart tablets to enhance both efficiency and accuracy.

We assembled a team of six experienced enumerators, all of whom were well-versed in the local language and cultural contexts. These enumerators underwent rigorous training, which included both classroom instruction and field-based exercises. Role-playing scenarios and detailed guidance on using the CAPI platform were key components of the training. Emphasis was placed on building rapport with respondents, navigating different interview scenarios, and adhering strictly to ethical guidelines. In-person interactions were pivotal, especially in regions where personal engagement is crucial to eliciting thorough and thoughtful responses.

The fieldwork spanned a six-month period, from October 2023 to March 2024. Throughout this time, research team members and field supervisors closely monitored the interview process, reviewed completed questionnaires for accuracy, and promptly addressed any issues. Data files were checked daily to ensure logical consistency, and discrepancies were

resolved through follow-up calls and additional checks. The final dataset underwent an extensive review to ensure its accuracy and representativeness. The CAPI methodology proved to be highly effective, enabling real-time monitoring, validation, and correction, all of which contributed to the high quality of the data collected. This approach, combining the reliability of electronic data capture with the personal engagement needed for accurate data collection, is particularly well-suited to the MENA region, addressing many of the challenges that traditional data collection methods face.

Preliminary summary statistics reveal a gender-balanced dataset, with a representative sample drawn from various regions across the country.

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	479	47.9	47.9	47.9
	Female	521	52.1	52.1	100.0
	Total	1000	100.0	100.0	

Age groups

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-24 years old	208	20.8	20.8	20.8
	25-34 years old	242	24.2	24.2	45.0
	35-44 years old	192	19.2	19.2	64.2
	45-54 years old	191	19.1	19.1	83.3
	55-64 years old	156	15.6	15.6	98.9

	65+ years old	11	1.1	1.1	100.0
	Total	1000	100.0	100.0	

Location

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Beirut (administrative)	70	7.0	7.0	7.0
	Beqaa	110	11.0	11.0	18.0
	Mount Lebanon	420	42.0	42.0	60.0
	Nabatieh	80	8.0	8.0	68.0
	North	200	20.0	20.0	88.0
	South	120	12.0	12.0	100.0
	Total	1000	100.0	100.0	